# ESSAYS ON TIME SERIES ECONOMETRICS

By

Cheol-Keun Cho

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2014

ABSTRACT

ESSAYS ON TIME SERIES ECONOMETRICS

By

Cheol-Keun Cho

Chapter 1 develops an asymptotic theory for testing the presence of structural change in a weakly dependent time series regression model. The cases of full structural change and partial structural change are considered. A HAC estimator is involved in the construction of the test statistics. Depending on how the long run variance for pre- and post-break regimes is estimated, two types of heteroskedasticity and autocorrelation (HAC) robust Wald statistics, denoted by $Wald^{(F)}$ and $Wald^{(S)}$ are analyzed. The fixed-$b$ asymptotics established by Kiefer and Vogelsang (2005) is applied to derive the limits of the statistics with the break date being treated as a priori. The fixed-$b$ limits turn out to depend on the location of break fraction and the bandwidth ratio as well as on the kernel being used. For both Wald statistics the limits capture the finite-sample randomness existing in the HAC estimators for the pre- and post-break regimes. The limit of $Wald^{(F)}$ further captures the finite-sample covariance between the pre-break estimators of regression parameters and the post-break estimators of the regression parameters. The fixed-$b$ limit stays the same and is pivotal for $Wald^{(F)}$ irregardless of whether some of the regressors are not subject to structural change. Critical values for the tests are obtained by simulation methods. Monte Carlo simulations compare the finite sample size properties of the two Wald statistics and a local power analysis is conducted to provide guidance on the power properties of the tests. This Chapter extends its analysis to cover the case of the break date being unknown. Supremum, mean and exponential Wald statistics are considered and finite sample size distortions are examined via simulations with newly tabulated fixed-$b$ critical values for these statistics.

Chapter 2 generalizes the structural change test developed in Chapter 1 while allowing

for a shift in the mean and(or) variance of the explanatory variable. Chapter 2 assumes the break date for the mean/variance is different from the possible break date for the regression parameters. The test is robust to serial correlation and heteroskedasticity of the error term and the explanatory variables. The fixed-$b$ theory is applied to derive the limits of the statistics. The asymptotic theory in this paper is based on a new set of high level conditions which incorporates the possibility of the moments shift and serves to provide pivotal limits of the test statistics.

Chapter 3 proposes a test of the null hypothesis of integer integration against the alternative of fractional integration. The null of integer integration is satisfied if the series is either $I(0)$ or $I(1)$, while the alternative is that it is $I(d)$ with $0 < d < 1$. The test is based on two statistics, the KPSS statistic and a new unit root test statistic . The null is rejected if the KPSS test rejects $I(0)$ and the unit root test rejects $I(1)$. The newly proposed unit root test is a lower-tailed KPSS test based on the first differences of the original data, so the test of the null of integer integration is called the "Double KPSS" test. Chapter 3 shows that the test has asymptotically correct size under the null that the series is either $I(0)$ or $I(1)$ and the test is consistent against $I(d)$ alternatives for all $d$ between zero and one. These statements are true under the assumption that the number of lags used in long-run variance estimation goes to infinity with sample size, but more slowly than sample size. Chapter 3 refers to this as "standard asymptotics." This requires some original asymptotic theory for the new unit root test, and also for the KPSS short memory test for the case that $d = 1/2$. Chapter 3 also considers "fixed-$b$ asymptotics" as in Kiefer and Vogelsang (2005). Finite-sample size and power of the Double KPSS test is investigated using both the critical values based on standard asymptotics and the critical values based on fixed-$b$ asymptotics. The new test is more accurate when it uses the fixed-$b$ critical values. The conclusion is that one can distinguish integer integration from fractional integration using the Double KPSS test, but it takes a rather large sample size to do so reliably.

To my beloved wife, Eun-Young,
and my daughter, Ellin

ACKNOWLEDGMENTS

Over the past five years in East Lansing, I have owed to many people for their support and help. Without them I couldn't get through the times I had. I want to deeply thank to Timothy J. Vogelsang, my advisor as well as coauthor, for his continuous support and excellent guidance. I can't simply imagine how I would be now without him. He is the most influential person for me at MSU. I learned so many things from him in teaching and research and a lot of meetings we had were among the best memorable moments for me. I will always miss those times with him after I leave East Lansing. I am also so grateful to Peter Schmidt, my another excellent advisor and Christine Amsler. They are both my coauthors. It was greatly pleasing and exciting to work with them. I remember how warm they welcomed their TA's including me at the appreciation dinner meeting every semester. Also, I am not forgetting how much they were helpful and considerate when I was in need of help. I could survive my second year with their support. Special thanks to Christine Amsler. I worked for her as research assistant but she was more than a supervisor and I deeply thank her again for encouraging me during those days. I also thank Leslie E. Papke, the graduate program director. She was so much responsive to what I was going through. I am also indebted to two other committee members, Jeffrey M. Wooldridge and Hira Koul at department of statistics and probability for what I have learned from their lectures and for their invaluable comments on my research. What I learned from them has become really important academic assets for me. I thank Margaret Lynch and Lori Jean Nichols for their support and tips in every important administration work. I had many good friends while I was in East Lansing. I could take a break from work and research and get refreshed while I was hanging out with them. They are Jaemin Baik, Mikayla Bowen, Sarah Brown, Hon Fong Cheah, Yeonjei Jung, Myoung-Jin Keay, Sang Hyun Kim, Soobin Kim, Sangjoon Lee, Seunghwa Rho and Sun Yu. I want to specially thank Reverend Borin Cho, Geum Jang, and many more at Lansing Korean

UMC. Speicial thanks also go to Sangryun Lee, Myungsook Kim, Bernhard G. Bodmann and Raul Susmel in Houston, Texas. I also deeply thank my parents, my parents in law and sisters and brother in Korea for their unceasing support and trust toward me. Last but not least I owe my deepest gratitude to my wife, Eun-Young Song and my daughter, Ellin for their love and for being together with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER 1

# Fixed-*b* Inference for Testing Structural Break in a Time Series Regresison

## 1.1 Introduction

Chapter 1 focuses on fixed-*b* inference of heteroskedasticity and autocorrelation (HAC) robust Wald statistics for testing for a structural break in a time series regression. Commonly used kernel-based nonparametric HAC estimators are considered to estimate the asymptotic variance. HAC estimators allow for arbitrary structure of the serial correlation and heteroskedasticity of weakly dependent time series and are consistent estimators of the long run variance under the assumption that the bandwidth ($M$) is growing at a certain rate slower than the sample size ($T$). Under this assumption, the Wald statistics converge to the usual chi-square distributions. However, because the critical values from the chi-square distribution are based on a consistency approximation for the HAC estimator, the chi-square limit does not reflect the often substantial finite sample randomness of the HAC estimator. Furthermore, the chi-square approximation does not capture the impact of the choice of the kernel or the bandwidth on the Wald statistics. The sensitivity of the statistics to the finite sample bias and variability of the HAC estimator is well known in the literature; Kiefer and Vogelsang (2005) among others have illustrated by simulation that the traditional inference with a HAC estimator can have poor finite sample properties.

Departing from the traditional approach, Kiefer and Vogelsang (2002a), Kiefer and Vogelsang (2002b) , and Kiefer and Vogelsang (2005) obtain an alternative asymptotic approximation by assuming that the ratio of the bandwidth to the sample size, $b = M/T$, is held constant as the sample size increases. Under this alternative nesting of the bandwidth, they obtain pivotal asymptotic distributions for the test statistics which depend on the choice of kernel and bandwidth tuning parameter. Simulation results indicate that the resulting fixed-$b$ approximation has less size distortions in finite samples than the traditional approach especially when the bandwidth is not small.

Theoretical explanations for the finite sample properties of the fixed-$b$ approach include the studies by Hashimzade and Vogelsang (2008), Jansson (2004), Sun, Phillips, and Jin (2008, hereafter SPJ), Gonçalves and Vogelsang (2011) and Sun (2013). Hashimzade and Vogelsang (2008) provide an explanation for the better performance of the fixed-$b$ asymptotics by analyzing the bias and variance of the HAC estimator. Gonçalves and Vogelsang (2011) provide a theoretical treatment of the asymptotic equivalence between the naive bootstrap distribution and the fixed-$b$ limit. Higher order theory is used by Jansson (2004), SPJ (2008) and Sun (2013) to show that the error in rejection probability using the fixed-$b$ approximation is more accurate than the traditional approximation. In a Gaussian location model Jansson (2004) proves that for the Bartlett kernel with bandwidth equal to sample size (i.e. $b = 1$), the error in rejection probability of fixed-$b$ inference is $O(T^{-1} \log T)$ which is smaller than the usual rate of $O(T^{-1/2})$. The results in SPJ (2008) complement Jansson's result by extending the analysis for a larger class of kernels and focusing on smaller values of bandwidth ratio $b$. In particular they find that the error in rejection probability of the fixed-$b$ approximation is $O(T^{-1})$ around $b = 0$. They also show that for positively autocorrelated series, which is typical for economic time series, the fixed-$b$ approximation has smaller error than the chi-square or standard normal approximation even when $b$ is assumed to decrease to zero although the stochastic orders are same.

In this chapter fixed-$b$ asymptotics is applied to testing for structural change in a weakly dependent time series regression. The structural change literature is now enormous and no attempt will be made here to summarize the relevant literature. Some key references include Andrews (1993) and Andrews and Ploberger (1994). Andrews (1993) treats the issue of testing for a structural break in the GMM framework when the one-time break date is unknown and Andrews and Ploberger (1994) derive asymptotically optimal tests. Bai and Perron (1998) consider multiple structural change occurring at unknown dates and cover the issues of estimation of break dates, testing for the presence of structural change and for the number of breaks. For a comprehensive survey of the recent structural break literature see Perron (2006).

Because a structural change in a regression relationship effectively divides the sample into two regimes, two different HAC estimators are considered when constructing Wald statistics. Asymptotically, estimators of the regression parameters are uncorrelated across the two regimes. One HAC estimator imposes this zero covariance restriction while the other estimator does not and this leads to two possible Wald statistics that can be used in practice denoted by $Wald^{(S)}$ and $Wald^{(F)}$. When the break date is known and coefficients of all regressors are subject to structural change (i.e. full structural change), both Wald statistics have pivotal fixed-$b$ limits but these limits are different. For the two Wald statistics the fixed-$b$ critical values increase as the bandwidth gets bigger and as the hypothesized break date is closer to the boundary of the sample. When some of the regressors are not subject to structural change (i.e. partial structural change), the Wald statistic based on the unrestricted HAC estimator ($Wald^{(F)}$) still has the same pivotal fixed-$b$ limit as in the case of full structural change.

A local power analysis is carried out under the fixed-$b$ approach and several patterns are reported. The local power of both Wald statistics is lower with bigger $b$ especially for the QS kernel. Power also improves as the break date gets closer to the middle of the sample regardless of bandwidth or kernel. With the within-regime effective bandwidths matched across the two statistics, the power difference is more evident when a big value of $b$ is used with the QS kernel.

A simulation study on the finite sample performance of fixed-$b$ inference reveals that the Wald statistic based on the unrestricted HAC estimator has better overall size performance especially when the quadratic spectral (QS) kernel and large bandwidths are used in the case of persistent data. In the comparison of the performance of fixed-$b$ inference with traditional inference, it is found that the latter is subject to substantially more size distortions. Similar to the known patterns in models without structural change, size distortions decrease as one uses bigger bandwidth or as the hypothesized break date is closer to the middle of the sample. When there is strong persistence in the data, rejections using fixed-$b$ critical values can be above nominal levels although these distortions are much smaller compared to traditional inference.

The remainder of this chapter is organized as follows. Section 1.2 reviews fixed-$b$ asymptotic theory in a regression with no structural change. In Section 1.3 the basic set up of the full structural-change model. Some preliminary results are laid out and the two HAC estimators are introduced. Section 1.4 derives the fixed-$b$ limits of the two Wald statistics. Section 1.5 explores patterns in the fixed-$b$ critical values when the break date is treated as a priori. Section 1.6 compares finite sample size across different approach of inferences, different choices of kernel and HAC estimators under various DGP specifications. Section 1.7 examines local power. Section 1.8 generalizes the results in Section 1.4 to a model with partial structural change. It is shown that the fixed-$b$ limit of $Wald^{(F)}$ in the partial structural change model is same as in the full structural change model. Section 1.9 covers the case where the break date is unknown providing the fixed-$b$ critical values. Section 1.10 summarizes and concludes. Proofs and supplemental results are collected in the Appendix.

## 1.2 Review of the Fixed-$b$ Asymptotics

Consider a weakly dependent time series regression with $p$ regressors given by

$$y_t = x_t'\beta + u_t. \tag{1.1}$$

Model (1.1) is estimated by ordinary least squares (OLS) giving

$$\widehat{\beta} = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{T} x_t y_t\right),$$

and $\widehat{u}_t = y_t - x_t'\widehat{\beta}$ are the OLS residuals. The centered OLS estimator is given by

$$\widehat{\beta} - \beta = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{T} v_t\right),$$

where $v_t = x_t u_t$. The asymptotic theory is based on the following two assumptions.

**Assumption 1.** $T^{-1}\sum_{t=1}^{[rT]} x_t x_t' \xrightarrow{p} rQ$, uniformly in $r \in [0,1]$, and $Q^{-1}$ exists.
**Assumption 2.** $T^{-1/2}\sum_{t=1}^{[rT]} x_t u_t = T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \Lambda W_p(r)$, $r \in [0,1]$, where $\Lambda\Lambda' = \Sigma$, and $W_p(r)$ is a $p \times 1$ standard Wiener process.

For a more detailed discussion about the regularity conditions under which Assumptions 1 and 2 hold, refer to Kiefer and Vogelsang (2002b). See Davidson (1994), Phillips and Durlauf (1986), Phillips and Solo (1992), and Wooldridge and White (1988) for more details. The matrix $Q$ is the non-centered variance-covariance matrix of $x_t$ and is typically estimated using the sample variance $\widehat{Q} = \frac{1}{T}\sum_{t=1}^{T} x_t x_t'$. The matrix $\Sigma \equiv \Lambda\Lambda'$ is the asymptotic variance of $T^{-1/2}\sum_{t=1}^{T} v_t$, which is, for a stationary series, given by

$$\Sigma = \Gamma_0 + \sum_{j=1}^{\infty}(\Gamma_j + \Gamma_j') \text{ with } \Gamma_j = E(v_t v_{t-j}').$$

Being a long run variance, $\Sigma$ is commonly estimated by the kernel-based nonparametric HAC estimator

$$\widehat{\Sigma} = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right) \widehat{v}_t \widehat{v}_s' = \widehat{\Gamma}_0 + \sum_{j=1}^{T-1} K\left(\frac{j}{M}\right) \left(\widehat{\Gamma}_j + \widehat{\Gamma}_j'\right),$$

where $\widehat{\Gamma}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{v}_t \widehat{v}_{t-j}'$, $\widehat{v}_t = x_t \widehat{u}_t$, $M$ is a bandwidth, and $K(\cdot)$ is a kernel ighting function.

Under some regularity conditions (see Andrews (1991), De Jong and Davidson (2000), Hansen (1992), Jansson (2002) or Newey and West (1987)), $\widehat{\Sigma}$ is a consistent estimator of $\Sigma$, i.e. $\widehat{\Sigma} \xrightarrow{p} \Sigma$. These regularity conditions include the assumption that $M/T \to 0$ as $M, T \to \infty$. This asymptotics is called 'traditional asymptotics' throughout this chapter.

In contrast to the traditional approach, fixed-$b$ asymptotics assumes $b = M/T$ is held constant as $T$ increases. Assumptions 1 and 2 are the only regularity conditions required to obtain a fixed-$b$ limit for $\widehat{\Sigma}$. Under the fixed-$b$ approach, for $b \in (0, 1]$, Kiefer and Vogelsang (2005) show that

$$\widehat{\Sigma} \Rightarrow \Lambda \mathbf{P}(b, \widetilde{W}_p) \Lambda', \tag{1.2}$$

where $\widetilde{W}_p(r) = W_p(r) - rW_p(1)$ is a $p$-vector of standard Brownian bridges and the form of the random matrix $\mathbf{P}(b, \widetilde{W}_p)$ depends on the kernel. Following Kiefer and Vogelsang (2005) three classes of kernels are considered. Let $H_p(r)$ denote a generic vector of stochastic processes. $H_p(r)'$ denotes its transpose. Then $\mathbf{P}(b, H_p)$ is defined as follows:

**case 1** If $K(x)$ is twice continuously differentiable everywhere (Class 1) such as the Quadratic Spectral kernel (QS), then

$$\mathbf{P}\left(b, H_p\right) \equiv -\int_0^1 \int_0^1 \frac{1}{b^2} K''\left(\frac{r-s}{b}\right) H_p(r) H_p(s)' drds, \tag{1.3}$$

where $K''(\cdot)$ is the second derivative of the kernel $K(\cdot)$.

**case 2** If $K(x)$ is the Bartlett kernel (Class 2), then

$$\mathbf{P}\left(b, H_p\right) \equiv \frac{2}{b}\int_0^1 H_p(r)H_p(r)'dr - \frac{1}{b}\int_0^{1-b}\left(H_p(r)H_p(r+b)' + H_p(r+b)H_p(r)'\right)dr. \tag{1.4}$$

**case 3** If $K(x)$ is continuous, $K(x) = 0$ for $|x| \geq 1$, and $K(x)$ is twice continuously differentiable everywhere except for $|x| = 1$ (Class 3) like Parzen kernel, then

$$\begin{aligned}\mathbf{P}\left(b, H_p\right) \equiv & -\int\int_{|r-s|<b} \frac{1}{b^2}K''\left(\frac{|r-s|}{b}\right) H_p(r)H_p(s)'drds \\ & + \frac{K'\_(1)}{b}\int_0^{1-b}\left(H_p(r+b)H_p(r)' + H_p(r)H_p(r+b)'\right)dr,\end{aligned} \tag{1.5}$$

where $K'\_(1) = \lim_{h\downarrow 0}\left[\left(K(1) - K(1-h)\right)/h\right]$, i.e., $K'\_(1)$ is the derivative of $K(x)$ from the left at $x = 1$.

Inference regarding $\beta$ is based on the asymptotic normality of the OLS estimator given by the result

$$\sqrt{T}(\widehat{\beta} - \beta) = \left(T^{-1}\sum_{t=1}^{T} x_t x_t'\right)^{-1}\left(T^{-1/2}\sum_{t=1}^{T} x_t u_t\right) \xrightarrow{d} Q^{-1}\Lambda W_p(1) \sim N(\mathbf{0}, Q^{-1}\Sigma Q^{-1}),$$

which follows directly from Assumptions 1 and 2. Suppose the null hypothesis of interest is, $H_0 : R_{l\times p}\beta = \mathbf{r}$, where $l$ is the number of the restrictions. Define the Wald statistic as

$$Wald = T\left(R\widehat{\beta} - \mathbf{r}\right)'\left(R\widehat{Q}^{-1}\widehat{\Sigma}\widehat{Q}^{-1}R'\right)^{-1}\left(R\widehat{\beta} - \mathbf{r}\right).$$

Under traditional asymptotics the well known result is obtained:

$$Wald \xrightarrow{d} \chi_l^2,$$

whereas under fixed-$b$ asymptotics,

$$Wald \Rightarrow W_l(1)' \left(\mathbf{P}(b, \widetilde{W}_l)\right)^{-1} W_l(1). \tag{1.6}$$

The limit of the Wald statistic depends on the sequence of bandwidths by which $\widehat{\Sigma}$ is indexed. It is important to note that we are not viewing the choice of sequence as a bandwidth rule for the choice of $M$ in practice. Rather, the point is that different asymptotic approximations are obtained for the two assumptions regarding $M$. Under the fixed-$b$ approach the random matrix $\mathbf{P}(b, \widetilde{W}_l)$ approximates the randomness in $\widehat{\Sigma}$ and its dependence on $M$ (through $b$) and the kernel $K(\cdot)$. In contrast the traditional approach approximates $\widehat{\Sigma}$ by a constant that does not depend on $M$ or the kernel.

Once structural change is allowed in the model, existing results in the fixed-$b$ literature no longer apply and new results are required.

## 1.3 Model of Structural Change and Preliminary Results

Consider a weakly dependent time series regression model with a structural break given by

$$y_t = w_t'\boldsymbol{\beta} + u_t, \tag{1.7}$$

$$w_t = \left(x_{1t}', x_{2t}'\right)', \ \boldsymbol{\beta} = \left(\beta_1', \beta_2'\right)',$$

$$x_{1t} = x_t \cdot \mathbf{1}(t \leq \lambda T), \ x_{2t} = x_t \cdot \mathbf{1}(t > \lambda T),$$

where $x_t$ is $p \times 1$ regressor vector, $\lambda \in (0,1)$ is a break point, and $\mathbf{1}(\cdot)$ is the indicator function. Let $[\lambda T]$ denote the integer part of $\lambda T$. Note that $x_{2t} = \mathbf{0}$ for $t = 1, 2, ..., [\lambda T]$ and $x_{1t} = \mathbf{0}$ for $t = [\lambda T] + 1, ..., T$. For the time being the potential break point $\lambda T$ is assumed to be known. The case of $\lambda$ being unknown is discussed in Section 1.9. The regression model (1.7) implies that coefficients of all explanatory variables are subject to potential structural change and this model is labeled the 'full' structural change model.

Of interest it is that the presence of a structural change in the regression parameters. Consider null hypothesis of the form

$$H_0 : R\boldsymbol{\beta} = \mathbf{0}, \tag{1.8}$$

where

$$\underset{(l\times 2p)}{R} = (R_1, \ -R_1),$$

and $R_1$ is an $l \times p$ matrix with $l \leq p$. Under the null hypothesis, it is being tested that one or more linear relationships on the regression parameter(s) do not experience structural change before and after the break point. Tests of the null hypothesis of no structural change about a subset of the slope parameters are special cases. For example one can test the null hypothesis that the slope parameter on the first regressor did not change by setting $R_1 = (1, 0, \ldots, 0)$. One can test the null hypothesis that none of the regression parameters have structural change by setting $R_1 = I_p$.

In order to establish the asymptotic limits of the HAC estimators and the Wald statistics, Assumptions 1 and 2 given in previous Section are sufficient. Those Assumptions imply that there is no heterogeneity in the regressors across the segments and the covariance structure of the errors are assumed to be same across segments as well.

**notation** For later use, define a $l \times l$ nonsingular matrix $A$ such that

$$R_1 Q^{-1} \Lambda \Lambda' Q^{-1} R_1' = AA', \tag{1.9}$$

and

$$R_1 Q^{-1} \Lambda W_p(r) \stackrel{d}{=} AW_l(r), \tag{1.10}$$

where $W_l(r)$ is $l \times 1$ standard Wiener process.

Focus on the OLS estimator of $\boldsymbol{\beta}$ given by

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{t=1}^{T} w_t w_t'\right)^{-1} \left(\sum_{t=1}^{T} w_t y_t\right),$$

which can be written for each segment as

$$\widehat{\beta}_1 = \left(\sum_{t=1}^{T} x_{1t} x_{1t}'\right)^{-1} \left(\sum_{t=1}^{T} x_{1t} y_t\right) = \left(\sum_{t=1}^{[\lambda T]} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{[\lambda T]} x_t y_t\right), \tag{1.11}$$

$$\widehat{\beta}_2 = \left(\sum_{t=1}^{T} x_{2t} x_{2t}'\right)^{-1} \left(\sum_{t=1}^{T} x_{2t} y_t\right) = \left(\sum_{t=[\lambda T]+1}^{T} x_t x_t'\right)^{-1} \left(\sum_{t=[\lambda T]+1}^{T} x_t y_t\right). \tag{1.12}$$

Fixed-$b$ results depend on the limiting behavior of the following partial sum process

$$\widehat{S}_t = \sum_{j=1}^{t} w_j \widehat{u}_j = \sum_{j=1}^{t} w_j \left(y_j - x_{1j}' \widehat{\beta}_1 - x_{2j}' \widehat{\beta}_2\right)$$

$$= \sum_{j=1}^{t} w_j \left(u_j - x_{1j}' \left(\widehat{\beta}_1 - \beta_1\right) - x_{2j}' \left(\widehat{\beta}_2 - \beta_2\right)\right). \tag{1.13}$$

Under Assumptions 1 and 2 the limiting behavior of $\widehat{\boldsymbol{\beta}}$ and the partial sum process $\widehat{S}_t$ are easily obtained.

**Proposition 1.** *Let $\lambda \in (0,1)$ be given. Suppose the data generation process is given by (1.7) and let $[rT]$ denote the integer part of $rT$ where $r \in [0,1]$. Then under Assumptions 1 and 2 as $T \to \infty$,*

$$\sqrt{T}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \begin{pmatrix} \sqrt{T}\left(\widehat{\beta}_1 - \beta_1\right) \\ \sqrt{T}\left(\widehat{\beta}_2 - \beta_2\right) \end{pmatrix} \xrightarrow{d} \begin{pmatrix} (\lambda Q)^{-1} \Lambda W_p(\lambda) \\ ((1-\lambda)Q)^{-1} \Lambda \left(W_p(1) - W_p(\lambda)\right) \end{pmatrix},$$

*and*

$$T^{-1/2} \widehat{S}_{[rT]} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \begin{pmatrix} F_p^{(1)}(r,\lambda) \\ F_p^{(2)}(r,\lambda) \end{pmatrix} \equiv \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} F_p(r,\lambda),$$

*where*

$$F_p^{(1)}(r,\lambda) = \left(W_p(r) - \frac{r}{\lambda}W_p(\lambda)\right)\cdot \mathbf{1}(r \le \lambda),$$

*and*

$$F_p^{(2)}(r,\lambda) = \left(W_p(r) - W_p(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W_p(1) - W_p(\lambda)\right)\right)\cdot \mathbf{1}(r > \lambda).$$

**Proof:** See the Appendix.

It is easily seen that the asymptotic distributions of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are Gaussian and are independent of each other. Hence the asymptotic covariance of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ is zero. The asymptotic variance of $\sqrt{T}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is given by $\mathbf{Q}_\lambda^{-1}\Omega\mathbf{Q}_\lambda^{-1}$, where

$$\mathbf{Q}_\lambda \equiv \begin{pmatrix} \lambda Q & \mathbf{0} \\ \mathbf{0} & (1-\lambda)Q \end{pmatrix} \text{ and } \Omega \equiv \begin{pmatrix} \lambda\Sigma & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\Sigma \end{pmatrix}.$$

In order to test the null hypothesis (1.8) HAC robust Wald statistisc are considered. These statistics are robust to heteroskedasticity and autocorrelation in the vector process, $v_t = x_t u_t$. The generic form of the robust Wald statistic is given by

$$Wald = T\left(R\widehat{\boldsymbol{\beta}}\right)'\left(R\widehat{\mathbf{Q}}_\lambda^{-1}\widehat{\Omega}\widehat{\mathbf{Q}}_\lambda^{-1}R'\right)^{-1}\left(R\widehat{\boldsymbol{\beta}}\right),$$

where

$$\widehat{\mathbf{Q}}_\lambda = \begin{pmatrix} T^{-1}\sum_{t=1}^{[\lambda T]} x_t x_t' & \mathbf{0} \\ \mathbf{0} & T^{-1}\sum_{t=[\lambda T]+1}^{T} x_t x_t' \end{pmatrix},$$

and $\widehat{\Omega}$ is an HAC robust estimator of $\Omega$.

Two estimators for $\widehat{\Omega}$ are analyzed. The first one, denoted by $\widehat{\Omega}^{(F)}$, is newly introduced in this chapter and it is constructed using the residuals directly from the dummy regression (1.7):

$$\widehat{\Omega}^{(F)} = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t\widehat{v}_s', \tag{1.14}$$

where $\widehat{v}_t = w_t\widehat{u}_t = \left(x_{1t}'\widehat{u}_t, \quad x_{2t}'\widehat{u}_t\right)'_{2p\times 1}$. Denote the components of $\widehat{v}_t$ as $\widehat{v}_t^{(1)} = x_{1t}\widehat{u}_t = x_t\widehat{u}_t\mathbf{1}(t \leq \lambda T)$ and $\widehat{v}_t^{(2)} = x_{2t}\widehat{u}_t = x_t\widehat{u}_t\mathbf{1}(t > \lambda T)$.

The other estimator is the HAC estimator appearing in existing literature (e.g. Bai and Perron (2006)) which is given by

$$\widehat{\Omega}^{(S)} = \begin{pmatrix} \lambda\widehat{\Sigma}^{(1)} & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\widehat{\Sigma}^{(2)} \end{pmatrix}, \tag{1.15}$$

where

$$\widehat{\Sigma}^{(1)} = \frac{1}{[\lambda T]}\sum_{t=1}^{[\lambda T]}\sum_{s=1}^{[\lambda T]} K\left(\frac{|t-s|}{M_1}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(1)\prime}, \tag{1.16}$$

$$\widehat{\Sigma}^{(2)} = \frac{1}{T-[\lambda T]}\sum_{t=[\lambda T]+1}^{T}\sum_{s=[\lambda T]+1}^{T} K\left(\frac{|t-s|}{M_2}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(2)\prime}. \tag{1.17}$$

Note that $\widehat{\Sigma}^{(1)}$ is constructed using $\widehat{v}_t^{(1)}$ (data from the pre-break regime) and uses the bandwidth $M_1$ and pre-break sample size $[\lambda T]$. Likewise $\widehat{\Sigma}^{(2)}$ is constructed using $\widehat{v}_t^{(2)}$ (data from the post-break regime) and uses the bandwidth $M_2$ and post-break sample size $T - [\lambda T]$.

In Andrews (1993) and Bai and Perron (2006), the estimator (1.15) is used to allow for a potential structural change in the long run variance $\Sigma$ itself. In this chapter the assumption is maintained that $\Sigma$ does not have structural change because allowing for heterogeneity in $\Sigma$ results in a non-pivotal fixed-$b$ limit of the Wald statistic. Finding an estimator of $\Omega$ that has a pivotal fixed-$b$ limit when $\Sigma$ has structural change is a topic of ongoing research.

At a superficial level, the estimators $\widehat{\Omega}^{(F)}$ and $\widehat{\Omega}^{(S)}$ look different but they are directly related. Using $\widehat{v}_t = (\widehat{v}_t^{(1)\prime}, \widehat{v}_t^{(2)\prime})^\prime$ one can write $\widehat{\Omega}^{(F)}$ as

$$\widehat{\Omega}^{(F)} = \begin{pmatrix} \widehat{\Omega}_{11}^{(F)} & \widehat{\Omega}_{12}^{(F)} \\ \widehat{\Omega}_{21}^{(F)} & \widehat{\Omega}_{22}^{(F)} \end{pmatrix}$$

$$= \begin{pmatrix} T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(1)\prime} & T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(2)\prime} \\ T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(1)\prime} & T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(2)\prime} \end{pmatrix}$$

$$= \begin{pmatrix} T^{-1}\sum_{t=1}^{[\lambda T]}\sum_{s=1}^{[\lambda T]} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(1)\prime} & T^{-1}\sum_{t=1}^{[\lambda T]}\sum_{s=[\lambda T]+1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(2)\prime} \\ T^{-1}\sum_{t=[\lambda T]+1}^{T}\sum_{s=1}^{[\lambda T]} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(1)\prime} & T^{-1}\sum_{t=[\lambda T]+1}^{T}\sum_{s=[\lambda T]+1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(2)\prime} \end{pmatrix} \tag{1.18}$$

$$= \begin{pmatrix} \lambda\widehat{\Sigma}^{(1)} & T^{-1}\sum_{t=1}^{[\lambda T]}\sum_{s=[\lambda T]+1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(2)\prime} \\ T^{-1}\sum_{t=[\lambda T]+1}^{T}\sum_{s=1}^{[\lambda T]} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(1)\prime} & (1-\lambda)\widehat{\Sigma}^{(2)} \end{pmatrix}$$

It is seen that the diagonal blocks of $\widehat{\Omega}^{(F)}$ are the same HAC estimators used by $\widehat{\Omega}^{(S)}$ except that the same bandwidth, $M$, is used for each diagonal block of $\widehat{\Omega}^{(F)}$ whereas the diagonal blocks of $\widehat{\Omega}^{(S)}$ can have different bandwidths. $\widehat{\Omega}^{(S)}$ is a restricted version of $\widehat{\Omega}^{(F)}$ with the off-diagonal blocks set to $\mathbf{0}$, i.e. $\widehat{\Omega}^{(S)}$ imposes a zero covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ matching the zero asymptotic covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$. In contrast $\widehat{\Omega}^{(F)}$ does not impose this zero covariance which is consistent with the possibility that the finite sample covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ is not equal to zero (which is true in general). Note however that $\widehat{\Omega}^{(F)}$ uses the same bandwidth for both regimes whereas $\widehat{\Omega}^{(S)}$ allows different bandwidths in the two regimes. Thus, from the bandwidth perspective, $\widehat{\Omega}^{(F)}$ is restrictive relative to $\widehat{\Omega}^{(S)}$.

The next Section provides asymptotic results for the two HAC robust Wald statistics under the traditional asymptotics and under the fixed-$b$ asymptotics.

## 1.4 Asymptotic Results

### 1.4.1 Asymptotic Limits under Traditional Approach

The goal of the traditional approach is to find conditions under which the HAC estimator is consistent. One requirement for consistency is that $M$ grows with the sample but at a slower rate. Then under additional regularity conditions, $\widehat{\Sigma}^{(1)}$ and $\widehat{\Sigma}^{(2)}$ are consistent estimators of $\Sigma$ and the limit of $\widehat{\Omega}^{(S)}$ is straightforwardly given by

$$\widehat{\Omega}^{(S)} = \underset{(2p\times 2p)}{\begin{pmatrix} \lambda\widehat{\Sigma}^{(1)} & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\widehat{\Sigma}^{(2)} \end{pmatrix}} \xrightarrow{p} \begin{pmatrix} \lambda\Sigma & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\Sigma \end{pmatrix}.$$

However establishing consistency of $\widehat{\Omega}^{(F)}$ requires some additional calculation beyond existing results in the literature.

**Proposition 2.** *Under regularity conditions for the consistency of the HAC estimators $\widehat{\Sigma}^{(1)}$ and $\widehat{\Sigma}^{(2)}$, as $T \to \infty$,*

$$\widehat{\Omega}^{(F)} \xrightarrow{p} \begin{pmatrix} \lambda\Sigma & \mathbf{0} \\ \mathbf{0} & (1-\lambda)\Sigma \end{pmatrix}_{2p\times 2p}.$$

**Proof:** See the Appendix.

Let $Wald^{(S)}$ denote the Wald statistic based on $\widehat{\Omega}^{(S)}$ and let $Wald^{(F)}$ denote the Wald statistic based on $\widehat{\Omega}^{(F)}$. Then the results given in this subsection, combined with Proposition 1, give us the limits of the test statistics under the traditional approach:

$$Wald^{(S)}, Wald^{(F)} \Rightarrow$$

$$\lambda(1-\lambda)\left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right)' \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right),$$

where $W_l$ is a $l \times 1$ standard Wiener process. Note that for any given value of $\lambda$ the limit follows a chi-square distribution with degrees of freedom $l$. While convenient, the chi-square limit does not capture the impact of the randomness of $\widehat{\Omega}^{(S)}$ and $\widehat{\Omega}^{(F)}$ on the Wald statistics in finite samples.

### 1.4.2 Asymptotic Limits under Fixed-$b$ Approach

Now fixed-$b$ limits for the HAC estimators and the test statistics are provided. The fixed-$b$ limits presented in next Lemma and Corollary approximate the diagonal blocks of $\widehat{\Omega}^{(F)}$ by random matrices. Also, it is shown that fixed-$b$ approach gives nonzero limit for the off-diagonal block, which further distinguishes fixed-$b$ asymptotics from the traditional asymptotics.

**Lemma 1.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1 and 2, as $T \to \infty$,*

$$\widehat{\Omega}^{(F)} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \times \mathbf{P}\left(b, F_p\left(r,\lambda\right)\right) \times \begin{pmatrix} \Lambda' & \mathbf{0} \\ \mathbf{0} & \Lambda' \end{pmatrix}, \tag{1.19}$$

*where*

$$F_p\left(r,\lambda\right) = \begin{pmatrix} F_p^{(1)}\left(r,\lambda\right) \\ F_p^{(2)}\left(r,\lambda\right) \end{pmatrix}, \tag{1.20}$$

$$F_p^{(1)}\left(r,\lambda\right) = \left(W_p(r) - \frac{r}{\lambda} W_p(\lambda)\right) \mathbf{1}\left(0 \le r \le \lambda\right), \tag{1.21}$$

$$F_p^{(2)}\left(r,\lambda\right) = \left(W_p(r) - W_p(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W_p(1) - W_p(\lambda)\right)\right) \mathbf{1}\left(\lambda < r \le 1\right), \tag{1.22}$$

*and $\mathbf{P}_p\left(b, F_p\left(r,\lambda\right)\right)$ is defined by (1.3), (1.4), and (1.5) with $H_p(r) = F_p\left(r,\lambda\right)$.*

**Proof:** See the Appendix.

Extra algebra leads to an alternative representation for $\mathbf{P}\left(b, F_p\left(r,\lambda\right)\right)$. The proof for this Corollary is omitted.

**Corollary 1.**

$$\mathbf{P}\left(b, F_p\left(r,\lambda\right)\right) = \begin{pmatrix} \mathbf{P}\left(b, F_p^{(1)}\left(r,\lambda\right)\right) & \mathbf{C}\left(b, F_p^{(1)}\left(r,\lambda\right), F_p^{(2)}\left(r,\lambda\right)\right) \\ \mathbf{C}\left(b, F_p^{(1)}\left(r,\lambda\right), F_p^{(2)}\left(r,\lambda\right)\right)' & \mathbf{P}\left(b, F_p^{(2)}\left(r,\lambda\right)\right) \end{pmatrix}, \quad (1.23)$$

*where*

$$\mathbf{C}\left(b, F_p^{(1)}\left(r,\lambda\right), F_p^{(2)}\left(r,\lambda\right)\right)$$

$$= \begin{cases} -\int_0^1 \int_0^1 \frac{1}{b^2} K''\left(\frac{|r-s|}{b}\right) F_p^{(1)}\left(r,\lambda\right) F_p^{(2)}\left(s,\lambda\right)' drds, \\ \frac{1}{b}\int_0^{1-b} F_p^{(1)}\left(r,\lambda\right) F_p^{(2)}\left(r+b,\lambda\right)' dr, \\ -\frac{\int\int_{|r-s|<b} K''\left(\frac{|r-s|}{b}\right) F_p^{(1)}(r,\lambda) F_p^{(2)}(s,\lambda)' drds}{b^2} + \frac{K'\_(1)\int_0^{1-b} F_p^{(1)}(r,\lambda) F_p^{(2)}(r+b,\lambda)' dr}{b}, \end{cases}$$

*for Class 1,2 and 3 kernels respectively.*

The expression for $\mathbf{P}\left(b, F_p\left(r,\lambda\right)\right)$ in this Corollary makes it easier to compare the fixed-$b$ limit of $\widehat{\Omega}^{(F)}$ with the standrad fixed-$b$ limit (see (1.2)) appearing in a non-structural change settings. Since each diagonal block of $\widehat{\Omega}^{(F)}$ is basically a HAC estimator (up to a scale factor; see (1.18)) based on one of the pre- or post- break data, its limit should take the same form as (1.2), which is verified in this Corollary. So each diagonal component of $\mathbf{P}\left(b, F_p\left(r,\lambda\right)\right)$ serves to reflect the randomness and bandwidth/kernel-dependence of the associated HAC estimator. Second, unlike the traditional approach, the fixed-$b$ limit of the off-diagonal component is nonzero. This implies the fixed-$b$ inference is able to take account of the covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ which is generally nonzero in finite samples.

Next Lemma contains the parallel result for $\widehat{\Omega}^{(S)}$.

**Lemma 2.** *Let $\lambda \in (0,1)$ and $b_1$, $b_2 \in (0,1]$ be given. Suppose $M_1 = b_1(\lambda T)$ and $M_2 = b_2(1-\lambda)T$. Then under Assumptions 1 and 2, as $T \to \infty$,*

$$\widehat{\Sigma}^{(1)} \Rightarrow \Lambda\left(\frac{\mathbf{1}}{\lambda}\mathbf{P}\left(b_1\lambda, F_p^{(1)}\left(r,\lambda\right)\right)\right)\Lambda', \ \ \widehat{\Sigma}^{(2)} \Rightarrow \Lambda\left(\frac{\mathbf{1}}{1-\lambda}\mathbf{P}\left(b_2\left(1-\lambda\right), F_p^{(2)}\left(r,\lambda\right)\right)\right)\Lambda',$$

*and*

$$\widehat{\Omega}^{(S)} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \times \begin{pmatrix} \mathbf{P}\left(b_1\lambda, F_p^{(1)}(r,\lambda)\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{P}\left(b_2(1-\lambda), F_p^{(2)}(r,\lambda)\right) \end{pmatrix} \times \begin{pmatrix} \Lambda' & \mathbf{0} \\ \mathbf{0} & \Lambda' \end{pmatrix},$$

**Proof:** See the Appendix.

The limits of the Wald statistics can be derived by using Lemmas 1 and 2 respectively, Theorem 1 delivers the result for $Wald^{(F)}$.

**Theorem 1.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1 and 2, as $T \to \infty$,*

$$Wald^{(F)} \Rightarrow \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right)'$$

$$\times \left(\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\right)^{-1} \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right) \tag{1.24}$$

**Proof:** See the Appendix.

The next Corollary provides an alternative representation for the limit given in (1.24).

**Corollary 2.** *For a given value of $\lambda \in (0,1)$, the fixed-b limit of $Wald^{(F)}$ has the same distribution as*

$$\frac{1}{\lambda(1-\lambda)}W_l(1)'\left(\frac{1}{\lambda}\mathbf{P}\left(\frac{b}{\lambda}, \widetilde{W}_l(r)\right) + \mathbf{CP}(\lambda, b)\right. \tag{1.25}$$

$$\left. + \frac{1}{1-\lambda}\mathbf{P}\left(\frac{b}{1-\lambda}, \widetilde{W}_l^*(r)\right) + \mathbf{CP}(\lambda, b)'\right)^{-1} W_l(1),$$

*where*

$$\mathbf{CP}\left(\lambda,b\right)=\begin{cases}\frac{\sqrt{\lambda}\sqrt{1-\lambda}\int_0^1\int_0^1 K''\left(\frac{|\lambda t-(1-\lambda)s-\lambda|}{b}\right)\widetilde{W}_l(t)\widetilde{W}_l^*(s)'dtds}{b^2} \text{ for Class-1 kernels,}\\ \frac{\int_0^{1-b}\widetilde{W}_l(\frac{r}{\lambda})\widetilde{W}_l^*\left(\frac{r+b-\lambda}{1-\lambda}\right)'1(\lambda-b<r\leq\lambda)dr}{b\sqrt{\lambda}\sqrt{1-\lambda}} \text{ for Class-2 kernels,}\\ \frac{\sqrt{\lambda}\sqrt{1-\lambda}\int_0^1\int_0^1 K''\left(\frac{|\lambda t-(1-\lambda)s-\lambda|}{b}\right)\widetilde{W}_l(t)\widetilde{W}_l^*(s)'1(|\lambda t-(1-\lambda)s-\lambda|<b)dtds}{b^2}\\ -\frac{\int_0^{1-b}K'_-(1)\widetilde{W}_l(\frac{r}{\lambda})\widetilde{W}_l^*\left(\frac{r+b-\lambda}{1-\lambda}\right)'1(\lambda-b<r\leq\lambda)dr}{b\sqrt{\lambda}\sqrt{1-\lambda}} \text{ for Class-3 kernels,}\end{cases}$$

*and $\widetilde{W}_l(r)$ and $\widetilde{W}_l^*(r)$ are $l\times 1$ Brownian Bridge processes which are independent of each other and of $W_l(1)$.*

**Proof:** See the Appendix.

The limit in (1.25) shows how the components of $\widehat{\Omega}^{(F)}$ affect the distribution of $Wald^{(F)}$. As mentioned earlier the random matrix $\mathbf{P}\left(\frac{b}{\lambda},\widetilde{W}_l(r)\right)$ reflects the random nature of $\widehat{\Omega}_{11}^{(F)}$ which is a part of asymptotic variance estimator of $\widehat{\beta}_1$. But as the first argument of $\mathbf{P}\left(\frac{b}{\lambda},\widetilde{W}_l(r)\right)$ manifests, this Corollary additionally reveals that the (effective) bandwidth for $\widehat{\Omega}_{11}^{(F)}$ turns out to be $\frac{b}{\lambda}$ not $b$. This picks up the fact that one implicitly uses the bandwidth ratio $\frac{b}{\lambda}$ for $\widehat{\Omega}_{11}^{(F)}$ when the researcher applies a bandwidth ratio $b$ for constructing $\widehat{\Omega}^{(F)}$. The second component $\mathbf{P}\left(\frac{b}{1-\lambda},\widetilde{W}_l^*(r)\right)$ is related with $\widehat{\Omega}_{22}^{(F)}$ (and $\widehat{\beta}_2$) in the exactly same fashion. Finally, having the third component $\mathbf{CP}\left(\lambda,b\right)$ as a part of the limit indicates that the fixed-$b$ inference captures the impact of finite-samples covariance on the variance of $\widehat{\beta}_1-\widehat{\beta}_2$ and on the inference for a structural change. The results for $Wald^{(S)}$ are available in the following Theorem and Corollary.

**Theorem 2.** *Let $\lambda\in(0,1)$ and $b_1$, $b_2\in(0,1]$ be given. Suppose $M_1=b_1\left(\lambda T\right)$ and $M_2=b_2(1-\lambda)T$. Then under Assumptions 1 and 2, as $T\to\infty$,*

$$Wald^{(S)}\Rightarrow\left(\frac{1}{\lambda}W_l(\lambda)-\frac{1}{1-\lambda}\left(W_l(1)-W_l(\lambda)\right)\right)'$$

$$\times\left(\frac{1}{\lambda}\mathbf{P}\left(b_1\lambda,F_l^{(1)}\left(r,\lambda\right)\right)+\frac{1}{1-\lambda}\mathbf{P}\left(b_2\left(1-\lambda\right),F_l^{(2)}\left(r,\lambda\right)\right)\right)^{-1} \quad (1.26)$$

$$\times\left(\frac{1}{\lambda}W_l(\lambda)-\frac{1}{1-\lambda}\left(W_l(1)-W_l(\lambda)\right)\right).$$

**Proof:** See the Appendix.

**Corollary 3.** *For a given value of $\lambda \in (0,1)$, the fixed-b limit of $Wald^{(S)}$ has the same distribution as*

$$\frac{1}{\lambda(1-\lambda)}W_l(1)'\left(\frac{1}{\lambda}\mathbf{P}\left(b_1,\widetilde{W}_l(r)\right)+\frac{1}{1-\lambda}\mathbf{P}\left(b_2,\widetilde{W}_l^*(r)\right)\right)^{-1}W_l(1). \tag{1.27}$$

**Proof:** See the Appendix.

The major difference between (1.27) and (1.25) lies in the term $\mathbf{CP}(\lambda,b)$ which is the limit associated with the cross product term of $\widehat{\Omega}^{(F)}$. The second difference is the bandwidth ratios applied to the HAC estimator. In constructing the HAC estimator $\widehat{\Omega}^{(F)}$, one can only choose a single value of bandwidth $M$ and accordingly a single value of bandwidth ratio $b \equiv \frac{M}{T}$. But this bandwidth $M$ implicitly determines the effective bandwidth ratios within each regimes: $\frac{M}{\lambda T} = \frac{bT}{\lambda T} = \frac{b}{\lambda}$ for the pre-break regime and $\frac{M}{(1-\lambda)T} = \frac{bT}{(1-\lambda)T} = \frac{b}{(1-\lambda)}$ for the post-break regime. On the other hand, when one uses $\widehat{\Omega}^{(S)}$, the researcher chooses two bandwidths $M_1$ and $M_2$ for each regime respectively and accordingly the two bandwidth ratios are set as $b_1 = \frac{M_1}{\lambda T}$ and $b_2 = \frac{M_2}{(1-\lambda)T}$. In order to accurately compare the test perfomance of $Wald^{(F)}$ and $Wald^{(S)}$, it should be ensured that the effective within-regime bandwidth ratios are the same across the two statistics. Specifically, if the bandwidth ratio used for $Wald^{(F)}$ is $b$, then one should pick bandwidth ratios as $b_1 = \frac{b}{\lambda}$ and $b_2 = \frac{b}{1-\lambda}$ for $Wald^{(S)}$. By doing this one can isolate the impact of the presence of $\mathbf{CP}(\lambda,b)$ on the inference.

## 1.5 Critical Values

The fixed-$b$ limiting distributions are nonstandard. Asymptotic critical values are easily obtained via simulations. The Wiener processes in the limiting distributions are approximated using scaled partial sums of 1,000 i.i.d. $N(0,1)$ random variables. Critical values are tabulated based on 50,000 replications. The 95% critical values for $l$=1 and 2 are provided for selected values of $b$ and $\lambda$ in Tables 1.1 through Table 1.13. The break fraction, $\lambda$, runs from 0.1 to 0.9 with 0.1 increments. The bandwidths considered are 0.02, 0.04, 0.06, 0.08, 0.1, 0.2, ..., 1. Tables 1.1 through 1.6 contain the critical values for $Wald^{(F)}$ and Table 1.7 through 1.13 provide the critical values for $Wald^{(S)}$. The critical value tables for $Wald^{(S)}$ only cover the case where $b_1 = b_2$ although Table 1.13 provides some critical values with $b_1 \neq b_2$.

Tables 1.1 through 1.6 display two main patterns of the critical values. First, for each given $\lambda$ the critical values increase as the bandwidth gets bigger. This is expected given the well known downward bias induced into HAC estimators from estimation error. The fixed-$b$ approximation captures this downward bias and reflects it through larger critical values. Second, for a given value of the bandwidth the critical values display a V-shaped pattern as a function of $\lambda$. As the break point moves closer to zero or one, the critical values increase and the minimum critical values occur at $\lambda = 0.5$. This V-shaped pattern is present but is less pronounced in the critical values for $Wald^{(S)}$.

## 1.6 Finite Sample Size

This Section reports the results of a finite sample simulation study that illustrates the performance of the fixed-$b$ critical values relative to the traditional critical values under the null hypothesis of no structural change. The data generating process (DGP) used in this simulation study is given by

$$
\begin{aligned}
y_t &= \beta_1 + \beta_2 x_t + u_t, \\
x_t &= \theta x_{t-1} + \epsilon_t, \\
u_t &= \rho u_{t-1} + \eta_t + \varphi \eta_{t-1},
\end{aligned}
$$

where $\epsilon_t$ and $\eta_t$ are independent of each other with $\epsilon_t, \eta_t \sim$ i.i.d. $N(0,1)$. The following parameter values are used:

$$
\begin{aligned}
(\beta_1, \beta_2) &= (0,\ 0), \\
\lambda &\in \{0.2,\ 0.4,\ 0.5\}, \\
\theta &\in \{0.5,\ 0.8,\ 0.9\}, \\
(\rho, \varphi) &\in \{(0,0),\ (0.5,0),\ (0.5,0.5),\ (0.9,\ 0.5),\ (0.9,\ 0.9)\},
\end{aligned}
$$

The DGP specifications for $(\theta,\ \rho,\ \varphi)$ include:

$$
\begin{aligned}
\text{A:}&\ \ \theta = 0.5,\ \rho = 0,\ \varphi = 0 \\
\text{B:}&\ \ \theta = 0.5,\ \rho = 0.5,\ \varphi = 0 \\
\text{C:}&\ \ \theta = 0.5,\ \rho = 0.5,\ \varphi = 0.5 \\
\text{D:}&\ \ \theta = 0.8,\ \rho = 0.5,\ \varphi = 0.5 \\
\text{E:}&\ \ \theta = 0.8,\ \rho = 0.9,\ \varphi = 0.5 \\
\text{F:}&\ \ \theta = 0.9,\ \rho = 0.9,\ \varphi = 0.9
\end{aligned}
$$

The value of $\theta$ measures the persistence of the time varying regressor $x_t$. The parameters $\rho$ and $\varphi$ jointly determine the serial correlation structure of the error term $u_t$. Bigger values of these three parameters leads to higher persistence of the time series $v_t \equiv x_t u_t$ except for specification A. Results for sample sizes $T = 50, 100, 500$ are reported and the number of replications is 2,500. The nominal level of all tests is 5%. $Wald^{(S)}$ and $Wald^{(F)}$ statistics are computed for testing the joint null hypothesis of no structural change in both the $\beta_1$ and $\beta_2$ parameters.

### 1.6.1 $Wald^{(S)}$ and $Wald^{(F)}$: Traditional Inference vs. Fixed-$b$ Inference

First the empirical null rejection probabilities of the two statistics are examined. The performance of the traditional chi-square critical values are compared with that of fixed-$b$ critical values. Tables 1.14 through 1.19 report empirical null rejection probabilities for $Wald^{(S)}$ test using the two critical values. In practice one can construct $\widehat{\Omega}^{(S)}$ using different bandwidths ratios $b_1$ and $b_2$ in the two regimes. To conserve space, only results for homogeneous bandwidths choices (i.e. $b_1 = b_2$) are provided. Tables 1.14 and 1.15 give results for DGP A for $\lambda = 0.2, 0.4$, Tables 1.16 and 1.17 give results for DGP D for $\lambda = 0.2, 0.4$, and Tables 1.18 and 1.19 give results for DGP E for $\lambda = 0.2, 0.4$. Each table reports results using the Bartlett and QS kernels for a range of values of $b$.

Consider the results for DGP A (no serial correlation in $v_t$) given in Tables 1.14 and 1.15. When $T = 50$, it is seen that both the Bartlett and QS kernel deliver tests that over-reject when the chi-square critical value is used. As the bandwidth gets bigger, this tendency to over-reject becomes more and more pronounced. Rejections using fixed-$b$ critical values are similar when small bandwidths are used. As the bandwidth increases, rejections using fixed-$b$ critical values systematically decrease towards the nominal level of 0.05. Increasing $T$ reduces over-rejections for both critical values when $b$ is small. In contrast, when $b$ is large, rejections using the chi-square critical value tend to persist as $T$ increases but rejections tends towards 0.05 when fixed-$b$ critical values are used. With $T = 500$, the QS kernel has empirical rejection probabilities close to 0.06 for all values of $b$ when fixed-$b$ critical values are used. It is clear that the fixed-$b$ approximation is often a substantial improvement over the traditional approximation. This is not surprising given that the fixed-$b$ approximation captures much of the randomness in the HAC estimator whereas the traditional approach treats the HAC estimator as a constant equal to its population value.

Tables 1.16-1.19 show the results for the data with stronger dependence. Patterns are similar to Tables 1.14-1.15 except that over-rejection problems tend to be higher in all cases. Using chi-square critical values leads to severe over-rejection problems regardless of bandwidth or kernel. Using fixed-$b$ critical values tends to reduce over-rejection problems especially when larger values of $b$ are used. The QS kernel tends to be much less over-sized than the Bartlett kernel. These finite sample patterns are very similar to patterns found by Kiefer and Vogelsang (2005) in non-structural change settings.

Comparing Tables 1.14, 1.16, 1.18 ($\lambda = 0.2$) with Tables 1.15, 1.17, 1.19 ($\lambda = 0.4$) it is seen that over-rejection problems tend to be greater with $\lambda = 0.2$ than with $\lambda = 0.4$. This makes sense because the $\lambda = 0.2$ case has a relatively small sample size for regime 1 compared to regime 2 whereas for $\lambda = 0.4$ the regime sample sizes are similar.

Next null rejection probabilities of the $Wald^{(F)}$ statistic are examined. In Figure 1.1 rejection probabilities for $Wald^{(F)}$ is plotted for the DGP given by $(\beta_1, \beta_2, \theta, \rho, \varphi) = (0,$ $0, 0.1, 0.5, 0.5)$ for the case of $T = 50$. Results are given for two kernels: Bartlett and QS and bandwidth ratio $b = 0.2$. For each kernel, rejections are computed using the chi-square critical value and the fixed-$b$ critical value. Rejections are plotted across a grid of break points given by $\lambda \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. Figure 1.1 shows the large size distortions associated with traditional inference. Over-rejections are substantial especially when $\lambda$ is close to the endpoints of the sample. In general a V-shaped pattern appears in over-rejections as a function of $\lambda$ with the least over-rejections occurring at the middle of the sample ($\lambda = 0.5$). Rejections using fixed-$b$ critical values are much closer to 0.05 and rejections are less sensitive to the location of the break point. Figure 1.1 also shows that the kernel matters. Rejections using the QS kernel when fixed-$b$ critical values are used are closer to 0.05.

Tables 1.20 through 1.25 provide additional results for $Wald^{(F)}$ for most of the DGPs given in Tables 1.14-1.19. Again rejections are reported using traditional chi-square critical values and fixed-$b$ critical values. The table also reports some rejections for $Wald^{(S)}$ which are discussed in the next subsection. The patterns in Tables 1.20-1.25 are generally very

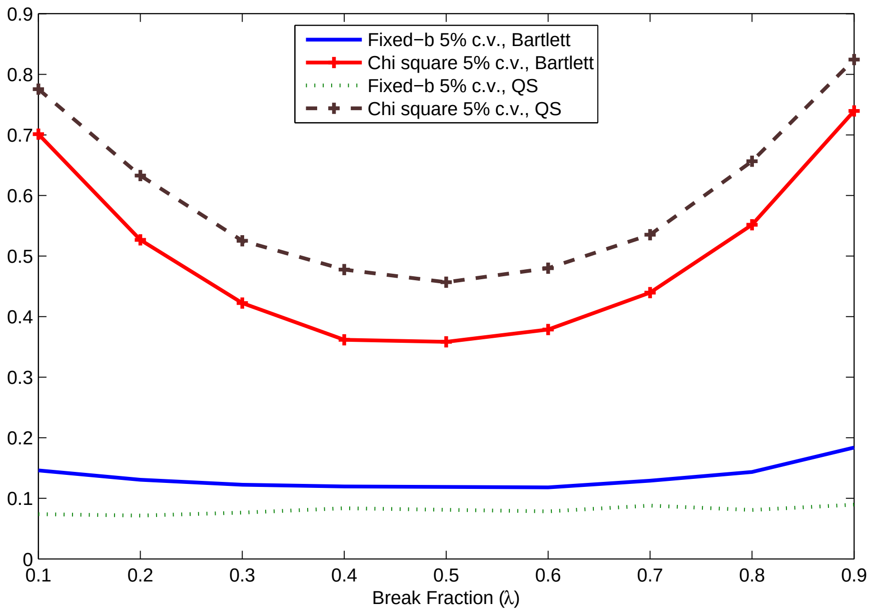


Figure 1.1: Empirical Null Rejection Probability of Structural Break Test using $Wald^{(F)}$
$\beta_1 = \beta_2 = 0$, $\theta = 0.1$, $\rho = 0.5$, $\phi = 0.5$, $b = 0.2$, $T = 50$, Replications=2,500

similar to the patterns seen in Tables 1.14-1.19. It is clear that the fixed-$b$ approximation is a substantial improvement over the traditional approximation.

### 1.6.2 Fixed-$b$ Inference: $Wald^{(S)}$ vs. $Wald^{(F)}$

In this subsection the finite sample rejections of the two Wald statistics are compared under the fixed-$b$ framework. As equation (1.25) and (1.27) manifest, the fixed-$b$ limits of the two statistics are different in two ways. First, the limit of $Wald^{(F)}$ has the extra component $\mathbf{CP}(\lambda, b)$. Second, the limit of $Wald^{(S)}$ depends on $b_1$ and $b_2$ which are the bandwidths for the two regimes respectively while the limit of $Wald^{(F)}$ depends on a single bandwidth $b$ implying effective bandwidths $\frac{b}{\lambda}$ and $\frac{b}{1-\lambda}$ in each regime. It is well known in the fixed-$b$ literature that larger bandwidths lead to less over-rejection problems when fixed-$b$ critical values are used. This serves to hedge against differences in effective bandwidths resulting in differences in over-rejections between $Wald^{(S)}$ and $Wald^{(F)}$ so that the impact of inclusion of the off-diagonal blocks used by $Wald^{(F)}$ can be isolated.

To keep bandwidth effects constant between $Wald^{(S)}$ and $Wald^{(F)}$, the bandwidths for $Wald^{(S)}$ are set as

$$b_1 = \frac{b}{\lambda} \text{ and } b_2 = \frac{b}{1-\lambda}. \tag{1.28}$$

where $b$ is the bandwidth for $Wald^{(F)}$. By using (1.28) it is ensured that the within regime bandwidth to (regime) sample size ratios for $Wald^{(S)}$ are the same as the bandwidth to (full) sample size ratio used by $Wald^{(F)}$.

The following eight bandwidths specifications satisfying (1.28) are used:

A′: $\lambda = 0.2, b = 0.04, b_1 = 0.2, b_2 = 0.05,$

B′: $\lambda = 0.2, b = 0.1, b_1 = 0.5, b_2 = 0.125,$

C′: $\lambda = 0.2, b = 0.2, b_1 = 1.0, b_2 = 0.25,$

D′: $\lambda = 0.5, b = 0.5, b_1 = 1.0, b_2 = 1.0,$

E′: $\lambda = 0.5, b = 1.0, b_1 = 2.0, b_2 = 2.0,$

F′: $\lambda = 0.4, b = 0.5, b_1 = 1.25, b_2 = 0.83,$

G′: $\lambda = 0.2, b = 0.5, b_1 = 2.5, b_2 = 0.625,$

H′: $\lambda = 0.2, b = 1.0, b_1 = 5.0, b_2 = 1.25.$

Tables 1.20 through 1.22 give results for relatively small bandwidths (A′, B′, and C′). With small bandwidths $Wald^{(S)}$ and $Wald^{(F)}$ have very similar size distortions. Tables 1.23 through 1.25 report results for the relatively large bandwidths (D′, E′, F′, G′, and H′). Note that compared to the small bandwidths case, there is noticeable difference in the performance of the two statistics specially when the QS kernel is used. As the DGP becomes more persistent (such as DGP F or G) the differences between $Wald^{(S)}$ and $Wald^{(F)}$ become more clear. Table 1.23 shows that under DGP G and bandwidth E′ the empirical null rejection probability is 19.08% for $Wald^{(S)}$ and 13.6% for $Wald^{(F)}$ when T=50 whereas in the Bartlett kernel case rejections are similar. The differences when T=100 are smaller

(15.96% vs. 11.92%); see Table 1.24. Once $T$ reaches 500 (Table 1.25), the differences between $Wald^{(S)}$ and $Wald^{(F)}$ are very small for either kernel. In general $Wald^{(F)}$ is less size distorted than $Wald^{(S)}$ when $T$ is relatively small and persistence in the data is not weak. For larger sample sizes the two tests have similar null rejection probabilities. From a size perspective, $Wald^{(F)}$ is preferred over $Wald^{(S)}$.

The better size performance of $Wald^{(F)}$ can be explained by examining $\mathbf{CP}(\lambda, b)$ in (1.25) and the corresponding finite sample counterparts $\widehat{\Omega}_{21}^{(F)}$ (or $\widehat{\Omega}_{12}^{(F)}$) in (1.18). $\widehat{\Omega}_{21}^{(F)}$ is estimating the sample covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ which is driven by finite sample covariance between $\widehat{v}_t^{(1)}$ and $\widehat{v}_t^{(2)}$ across the two regimes. Larger bandwidths lead to $\widehat{\Omega}_{21}^{(F)}$ and $\widehat{\Omega}_{12}^{(F)}$ estimators that better reflect the finite sample covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ and $\mathbf{CP}(\lambda, b)$ captures the impact of $\widehat{\Omega}_{21}^{(F)}$ and $\widehat{\Omega}_{12}^{(F)}$ on the sampling distribution of $Wald^{(F)}$. As the data becomes more persistent, the covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ becomes more pronounced in small samples so there is benefit to including $\widehat{\Omega}_{21}^{(F)}$ and $\widehat{\Omega}_{12}^{(F)}$ in $\widehat{\Omega}$. As $T$ increases, the covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ and it becomes more reasonable to impose the restriction that the off-diagonal blocks of $\widehat{\Omega}$ are zero as done by $Wald^{(S)}$.

## 1.7 Local Power Analysis of Fixed-$b$ Inference

This Section investigates the local power of $Wald^{(S)}$ and $Wald^{(F)}$ under the fixed-$b$ approach. The local alternative is given by

$$H_1 : R\beta = T^{-1/2} \underset{(l\times 1)}{\delta^*} \tag{1.29}$$

with $R = (R_1, \; -R_1)$. Under the local alternative (1.29), the fixed-$b$ limits of the Wald statistics are easily obtained as

$$\begin{aligned} Wald^{(F)} \Rightarrow & \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right) - A^{-1}\delta^*\right)' \\ & \times \left(\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\right)^{-1} \\ & \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right) - A^{-1}\delta^*\right),. \end{aligned}$$

$$\begin{aligned} Wald^{(S)} \Rightarrow & \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right) - A^{-1}\delta^*\right)' \\ & \times \left(\frac{1}{\lambda}\mathbf{P}\left(b_1\lambda, F_l^{(1)}(r,\lambda)\right) + \frac{1}{1-\lambda}\mathbf{P}\left(b_2(1-\lambda), F_l^{(2)}(r,\lambda)\right)\right)^{-1} \\ & \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right) - A^{-1}\delta^*\right), \end{aligned}$$

where the nonsingular matrix $A$ is defined in equation (1.10).

### 1.7.1 Comparison of the Local Asymptotic Power of $Wald^{(S)}$ and $Wald^{(F)}$

In this subsection the local asymptotic power of $Wald^{(S)}$ and $Wald^{(F)}$ are compared for the case $R = (I_2, \; -I_2)$. The bandwidth specifications are the same as those used in Section 1.6.2. Local asymptotic power was computed using the same simulation methods used to compute asymptotic fixed-$b$ critical values with $A^{-1}\delta^* = (\delta, \delta)'$, where $\delta$ is a scalar parameter. Figures 1.2-1.5 plot local asymptotic power for the Bartlett and QS kernels for a selection breakpoint and bandwidth specifications. As the figures illustrate, there is no substantial difference in the local asymptotic power between $Wald^{(S)}$ and $Wald^{(F)}$ for the case of the Bartlett kernel. However, when QS kernel is used, differences in power emerge with large bandwidths as depicted by Figures 1.3-1.5. For example, using $b = 1$ as shown in Figures 1.4 and 1.5, one can see substantial power loss associated with $Wald^{(F)}$ with the QS kernel. Comparing the Bartlett and QS kernels, the Bartlett kernel delivers higher

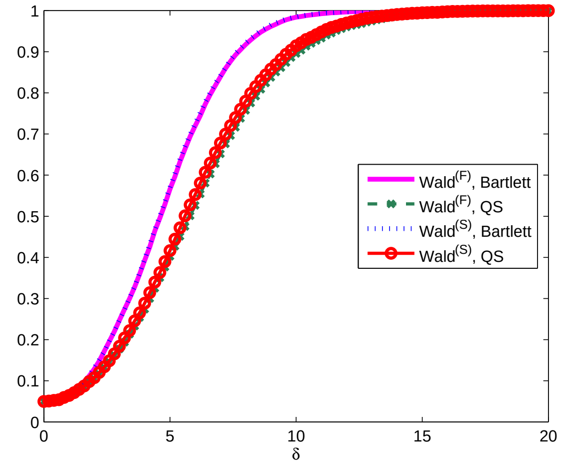


Figure 1.2: Local Power, Bartlett and QS,
$\lambda = 0.2$, $b = 0.2$, $b_1 = 1$, $b_2 = 0.25$

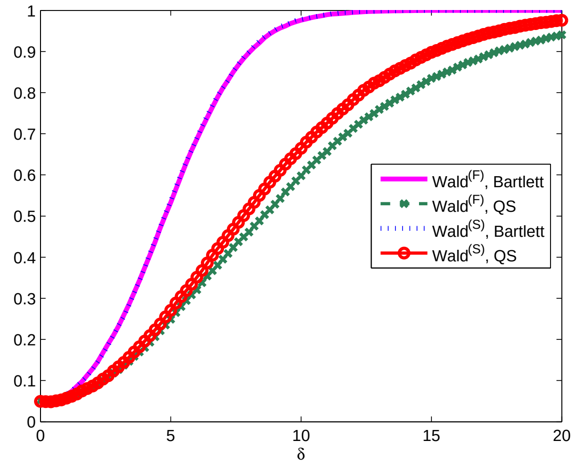


Figure 1.3: Local Power, Bartlett and QS,
$\lambda = 0.2$, $b = 0.5$, $b_1 = 2.5$, $b_2 = 0.625$

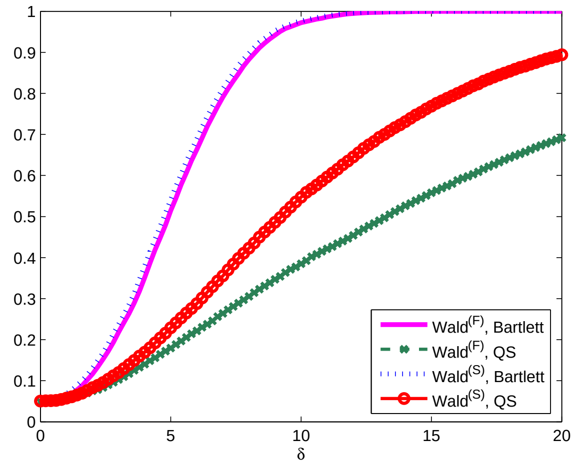


Figure 1.4: Local Power, Bartlett and QS,
$\lambda = 0.2$, $b = 1$, $b_1 = 5$, $b_2 = 1.25$

power than the QS kernel for both Wald statistics. The loss of power associated with $Wald^{(F)}$ relative to $Wald^{(S)}$ occurs exactly for the kernel and bandwidth specifications that resulted in $Wald^{(F)}$ having less size distortions than $Wald^{(S)}$. Similarly, the higher power of the Bartlett kernel relative to the QS kernel comes at the cost of greater size distortions. As has been documented repeatedly in the fixed-$b$ literature, there is an inherent trade-off between reduction of over-rejection problems and loss of power. Bandwidth/kernel combinations that reduce over-rejection problems also reduce power.

### 1.7.2 Impact of Breakpoint Location and Bandwidth on Power

The next focus is on the impact of the breakpoint location, $\lambda$, and bandwidths on local asymptotic power. Figures 1.6-1.9 display local asymptotic power for a range of $\lambda$. Each figure depicts power curves for the range of $\lambda = 0.1, 0.2, 0.3, 0.4, 0.5$ for the bandwidth ratio $b = 0.1$. Figures 1.6 and 1.7 depict power for the $Wald^{(F)}$ statistic for the Bartlett and QS kernels respectively. Figures 1.8 and 1.9 depict power for $Wald^{(S)}$ for the Bartlett and QS kernels respectively. For a given $\lambda$, the bandwidths $b_1$ and $b_2$ for $Wald^{(S)}$ are chosen so that they match the within-regime effective bandwidths of $Wald^{(F)}$, $\frac{b}{\lambda}$ and $\frac{b}{1-\lambda}$. For a given $b$, smaller values of $\lambda$ implies larger values of $\max\left(\frac{b}{\lambda}, \frac{b}{1-\lambda}\right)$ and the local

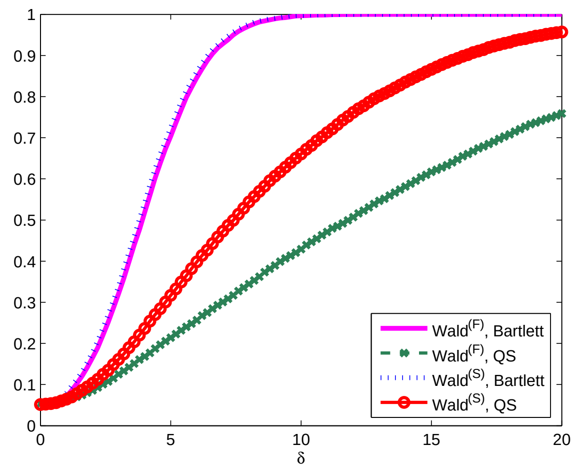


Figure 1.5: Local Power, Bartlett and QS,
$\lambda = 0.5$, $b = 1$, $b_1 = 2$, $b_2 = 1$

power for $Wald^{(F)}$ decreases as $\max\left(\frac{b}{\lambda}, \frac{b}{1-\lambda}\right)$ increases. First notice that regardless of bandwidth or kernel, power is lowest for $\lambda = 0.1$ and steadily rises as $\lambda$ increases to 0.5. It should be noted that power is symmetric in $\lambda$ and power for values of $\lambda$ larger than 0.5 can be inferred from the figures. Second, comparing Figures 1.6 and 1.8 (Bartlett) with Figures 1.7 and 1.9 (QS) the Bartlett kernel gives higher local power than the QS kernel especially for large within-regime effective bandwidths. A similar finding was documented by Kiefer and Vogelsang (2005) in models without structural change.

To isolate the impact of bandwidths on power, Figures 1.10 through 1.13 depict power for a range of $b$ values for a given kernel and $\lambda = 0.5$. These figures illustrate that increasing the bandwidth generally reduces power and this is especially true for the QS kernel. For the Bartlett kernel, once $b = 0.5$, further increases in $b$ have little effect on power. In contrast, increasing $b$ from 0.5 to 1.0 significantly reduces power when using the $QS$ kernel although the drop in power is even more dramatic when $b$ increases from 0.1 to 0.5. That power of $Wald^{(F)}$ and $Wald^{(S)}$ is decreasing in the bandwidth is not surprising given that this pattern holds in models without structural change.

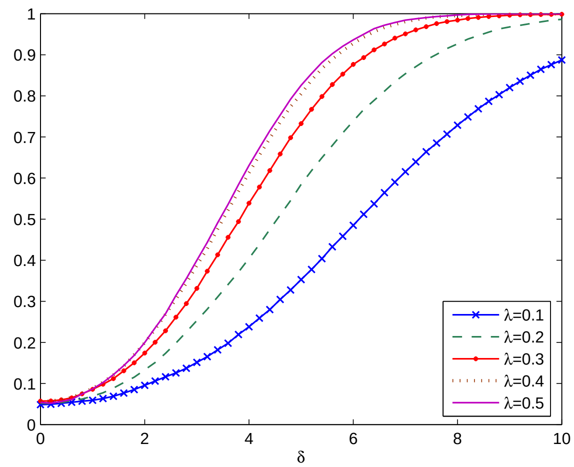


Figure 1.6: Local Power, $Wald^{(F)}$, Bartlett, $b = 0.1$

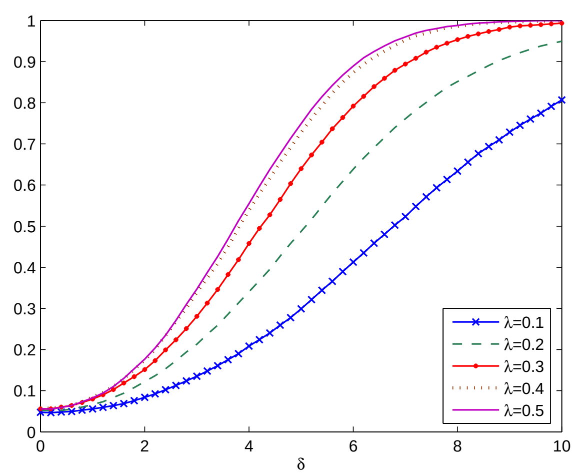


Figure 1.7: Local Power, $Wald^{(F)}$, QS, $b = 0.1$

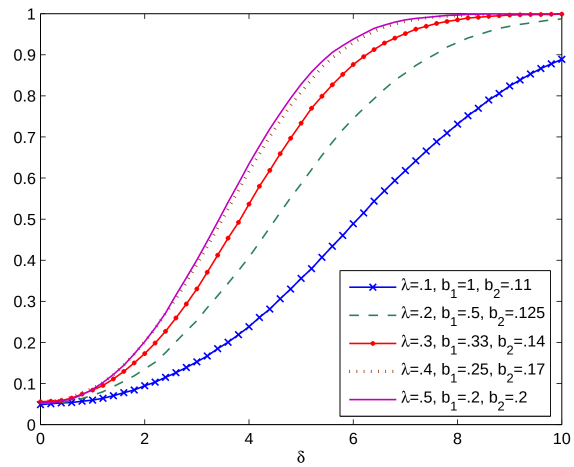


Figure 1.8: Local Power, $Wald^{(S)}$, Bartlett

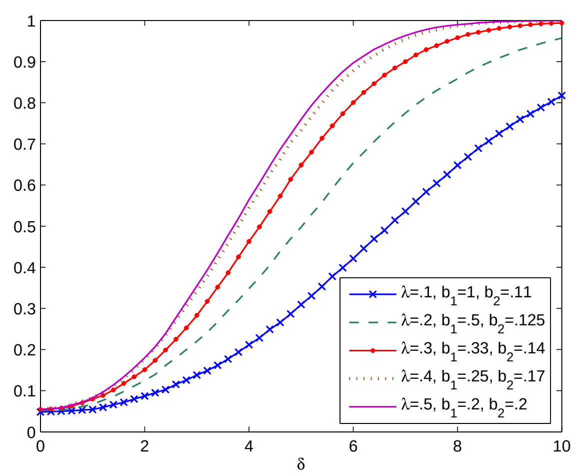


Figure 1.9: Local Power, $Wald^{(S)}$, QS

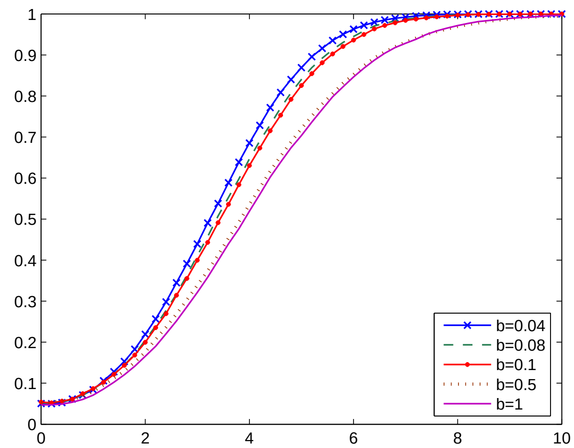


Figure 1.10: Local Power, $Wald^{(F)}$, Bartlett, $\lambda = 0.5$

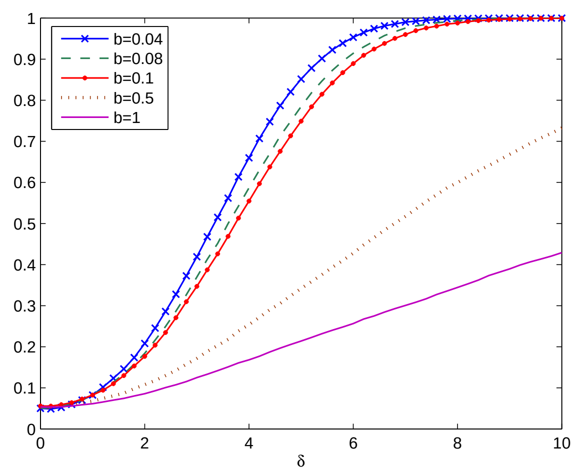


Figure 1.11: Local Power, $Wald^{(F)}$, QS, $\lambda = 0.5$

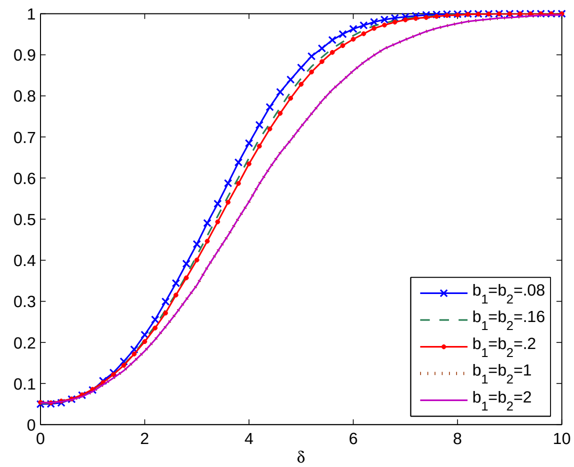


Figure 1.12: Local Power, $Wald^{(S)}$, Bartlett, $\lambda = 0.5$

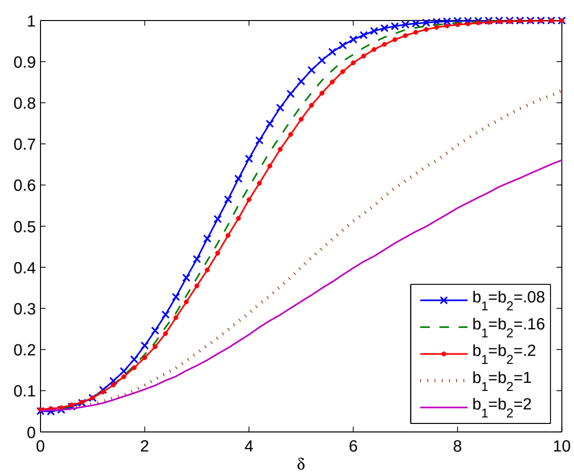


Figure 1.13: Local Power, $Wald^{(S)}$, QS, $\lambda = 0.5$

## 1.8 Partial Structural Change Model

This Section derives the fixed-$b$ limit of $Wald^{(F)}$ in the partial structural change model. The main result of this Section is that the limit is the same as the limit for the full structural change model. In contrast, the $Wald^{(S)}$ statistic in the partial structural change case has a non-pivotal fixed-$b$ limit. A modificaiton of the $Wald^{(S)}$ statistic that has a pivotal fixed-$b$ limit is briefly discussed.

### 1.8.1 Setup and Assumptions

The regression model with partial structural change is given by

$$\begin{aligned} y_t &= z_t'\alpha + x_{1t}'\beta_1 + x_{2t}'\beta_2 + u_t \\ &= z_t'\alpha + X_t'\boldsymbol{\beta} + u_t, \end{aligned} \tag{1.30}$$

where $x_t$ is $p \times 1$ and $z_t$ is $q \times 1$ vector and

$$\begin{aligned} &x_{1t} = x_t 1(t \leq \lambda T),\ x_{2t} = x_t 1(t > \lambda T), \\ &X_t = (x_{1t}'\ x_{2t}')',\ \text{and}\ \boldsymbol{\beta} = (\beta_1'\ \beta_2')'. \end{aligned}$$

The coefficients on the $x_t$ regressors are unrestricted in terms of a structural change whereas the coefficients on the $z_t$ regressors are assumed to not have structural change. Denote

$$\begin{aligned} &y = (y_1, y_2, \ldots, y_T)',\ X = (X_1, X_2, \ldots X_T)', \\ &Z = (z_1, z_2, \ldots, z_T)',\ u = (u_1, u_2, \ldots, u_T)'. \end{aligned}$$

The parameters $(\alpha, \boldsymbol{\beta})$ are estimated by OLS and the OLS residual vector can be written as

$$\widehat{u} = \widetilde{y} - \widetilde{X}\widehat{\boldsymbol{\beta}} = u - \widetilde{X}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) - P_Z u,$$

where

$$\widetilde{y} = (I - P_Z)\, y,\ \widetilde{X} = (I - P_Z)\, X, \text{ and } P_Z = Z(Z'Z)^{-1}Z'.$$

The residual for a individual observation is given by

$$\widehat{u}_t = u_t - \widetilde{X}_t'\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) - z_t'\left(Z'Z\right)^{-1} Z'u. \tag{1.31}$$

Also, note that

$$\widetilde{X}_t = X_t - X'Z(Z'Z)^{-1}z_t = \begin{pmatrix} \underset{p\times 1}{\widetilde{X}_t^{(1)}} \\ \underset{p\times 1}{\widetilde{X}_t^{(2)}} \end{pmatrix}.$$

The following assumptions replace Assumptions 1 and 2 in Section 1.2:

**Assumption 1′.** $T^{-1/2}\sum_{t=1}^{[rT]}\begin{pmatrix} x_t u_t \\ z_t u_t \end{pmatrix} \Rightarrow \Lambda W_{p+q}(r) \equiv \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix} W_{p+q}(r)$, where $\Lambda_1$ is a $p \times (p+q)$ matrix, $\Lambda_2$ is a $q \times (p+q)$ matrix, and $W_{p+q}(r)$ is a $(p+q) \times 1$ vector of independent Wiener process.

**Assumption 2′.** $p\lim \frac{1}{T}\sum_{t=1}^{[rT]} z_t z_t' = rQ_{ZZ}$, $p\lim \frac{1}{T}\sum_{t=1}^{[rT]} x_t x_t' = rQ_{xx}$, and $p\lim \frac{1}{T}\sum_{t=1}^{[rT]} x_t z_t' = rQ_{xZ}$ uniformly in $r \in [0,1]$, and there exist $Q_{ZZ}^{-1}$ and $Q_{xx}^{-1}$.

### 1.8.2 Asymptotic Limits

Continue to focus on tests of the null hypothesis of no structural change in the $x_t$ slope parameters of the form

$$H_0 : R\boldsymbol{\beta} = \mathbf{r}$$

with

$$\underset{l\times 2p}{R} = \begin{pmatrix} \underset{l\times p}{R_1}, & \underset{l\times p}{-R_1} \end{pmatrix} \text{ and } \mathbf{r} = \mathbf{0}. \tag{1.32}$$

Recall that the OLS estimator, $\widehat{\boldsymbol{\beta}} = \left(\widehat{\beta}_1', \widehat{\beta}_2'\right)'$ can be rewritten as

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{t=1}^{T} \widetilde{X}_t \widetilde{X}_t'\right)^{-1} \left(\sum_{t=1}^{T} \widetilde{X}_t \widetilde{y}_t\right). \tag{1.33}$$

**Proposition 3.** *Under Assumptions 1′ and 2′, as $T \to \infty$*

$$T^{1/2}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \xrightarrow{d} Q_{\widetilde{X}\widetilde{X}}^{-1} \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \\ \Lambda_1 \left(W_{p+q}(1) - W_{p+q}(\lambda)\right) - (1-\lambda) Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \end{pmatrix},$$

*and*

$$\sqrt{T}\left(R\widehat{\boldsymbol{\beta}} - \mathbf{r}\right) \overset{H_0}{\Rightarrow} R_1 Q_{xx}^{-1} \Lambda_1 \left(\frac{1}{\lambda} W_{p+q}(\lambda) - \frac{1}{1-\lambda}\left(W_{p+q}(1) - W_{p+q}(\lambda)\right)\right), \tag{1.34}$$

*where* $Q_{\widetilde{X}\widetilde{X}} = p\lim\left(T^{-1}\sum_{t=1}^{T} \widetilde{X}_t \widetilde{X}_t'\right)$.

**Proof:** See the Appendix.

As seen from the above proposition, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are not asymptotically independent in the partial structural change regression model. This is true because we are projecting out the variation of explanatory variables $z_t$ so that $\widehat{\beta}_1$ and $\widehat{\beta}_2$ depend on the entire series of $x_t$ and $z_t$. The dichotomy that $\widehat{\beta}_1$ is dependent only on the pre-break data and $\widehat{\beta}_2$ depends only on the post-break data no longer holds in the partial structural change model. The dependence manifests in the common term, $Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1)$ in the above Proposition. However, this term cancels out in (1.34) when the restriction matrix takes the form of (1.32). As a result, and also as suggested by equation (1.34), one only needs to estimate $\Lambda_1\Lambda_1'$ for the inference on structural change. But, $\Lambda_1\Lambda_1'$ can be easily estimated using $x_{1t}\widehat{u}_t$ or $x_{2t}\widehat{u}_t$ and the corresponding version of $Wald^{(S)}$ can be defined by constructing

$\widehat{\Sigma}^{(1)}$ with $x_{1t}\widehat{u}_t$ and constructing $\widehat{\Sigma}^{(2)}$ with $x_{2t}\widehat{u}_t$. By doing this, we are ignoring the need to project $z_t$ out of the $x_t$ variables but this ignorance does not make the inference invalid. The next question is whether in conducting inference one can still take into account the dependence between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ which is not due to the presence of $z_t$. This can be verified by finding a version of $Wald^{(F)}$ which has a pivotal limit in the partial structural change model. The answer is positive as below.

The Wald statistic is given by

$$Wald = T\left(R\widehat{\boldsymbol{\beta}}\right)'\left(R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}\widehat{\Omega}\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}R'\right)^{-1}\left(R\widehat{\boldsymbol{\beta}}\right), \tag{1.35}$$

where $\widehat{Q}_{\widetilde{X}\widetilde{X}} = T^{-1}\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'$. For constructing $Wald^{(F)}$, consider a HAC estimator $\widehat{\Omega}^{(F)}$ which is computed using $\left\{\widetilde{X}_t\widehat{u}_t\right\}_{t=1}^{T}$:

$$\widehat{\Omega}^{(F)} = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}K\left(\frac{|t-s|}{M}\right)\widehat{\xi}_t\widehat{\xi}_s', \tag{1.36}$$

where $\widehat{\xi}_t = \widetilde{X}_t\widehat{u}_t$. This is a straightforward extension of $Wald^{(F)}$ to the case of partial structural change.

Next Lemma provides the limit of the the scaled partial sum process of $\widehat{\xi}_t$ premultiplied by an appropriate term.

**Lemma 3.** *Let $\widehat{S}_t^{\xi} = \sum_{j=1}^{t}\widehat{\xi}_j$. Under Assumptions 1′ and 2′, as $T \to \infty$,*

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1/2}\widehat{S}_{[rT]}^{\xi} \Rightarrow R_1Q_{xx}^{-1}\Lambda_1\left(\frac{1}{\lambda}F_{p+q}^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_{p+q}^{(2)}(r,\lambda)\right),$$

*where*

$$F_{p+q}^{(1)}(r,\lambda) = \left(W_{p+q}(r) - \frac{r}{\lambda}W_{p+q}(\lambda)\right)\mathbf{1}(0 \le r \le \lambda),$$

$$F_{p+q}^{(2)}(r,\lambda) = \left(W_{p+q}(r) - W_{p+q}(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W_{p+q}(1) - W_{p+q}(\lambda)\right)\right)\mathbf{1}(\lambda < r \le 1).$$

**Proof:** See the Appendix.

As Lemma 3 shows, the partial sums of the inputs to $\widehat{\Omega}^{(F)}$ are asymptotically proportional to the same nuisance parameters as $\sqrt{T}\left(R\widehat{\boldsymbol{\beta}}-\mathbf{r}\right)$. This is the key condition for an asymptotic pivotal fixed-$b$ limit. The next Theorem provides the fixed-$b$ limit of $Wald^{(F)}$.

**Theorem 3.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1′ and 2′, $Wald^{(F)}$ weakly converges to the same limit in (1.24), i.e. as $T \to \infty$,*

$$Wald^{(F)} \Rightarrow \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right)'$$
$$\times \left(\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\right)^{-1} \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right).$$

**Proof:** See the Appendix.

According to Theorem 3, the limit of $Wald^{(F)}$ in the partial structural change model is the same as in the full structural change model.

Getting back to $Wald^{(S)}$, as mentioned earlier, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ are no longer asymptotically uncorrelated in the partial structural change model. Therefore, forcing the covariance between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ to be zero cannot be justified any longer. Even though forcing the covariance to be zero is not theoretically justified, even asymptotically, $Wald^{(S)}$ can be modified so that it has a pivotal fixed-$b$ limit. This is obtained by using $\widehat{v}_t^{(1)} = x_{1t}\widehat{u}_t = x_t\widehat{u}_t\mathbf{1}(t \leq \lambda T)$ for constructing $\widehat{\Sigma}^{(1)}$ and $\widehat{v}_t^{(2)} = x_{2t}\widehat{u}_t = x_t\widehat{u}_t\mathbf{1}(t > \lambda T)$ for $\widehat{\Sigma}^{(2)}$ with $\widehat{u}_t$ being defined in (1.31). One can easily show that the limit of this $Wald^{(S)}$ is the same as in (1.26). It is tempted to use $\widetilde{X}_t\widehat{u}_t\mathbf{1}(t \leq \lambda T)$ for constructing $\widehat{\Sigma}^{(1)}$ (and similarly for $\widehat{\Sigma}^{(2)}$ using $\widetilde{X}_t\widehat{u}_t\mathbf{1}(t > \lambda T)$). But this version of $Wald^{(S)}$ turns out to have a non-pivotal fixed-$b$ limit. This can be easily verified by deriving the limits of the associated partial sum processes.

## 1.9 When the Break Date is Unknown

Tests for a potential structural break with a unknown break date are well studied in Andrews (1993) and Andrews and Ploberger (1994). Andrews (1993) considers several tests based on the supremums across break points of Wald and LM statistics and shows they are asymptotically equivalent. Andrews and Ploberger (1994) derive tests that maximize average power across potential breakpoints. $Wald^{(F)}$ statistic is the only focus of this Section. Given a value of $b$, the test statistic is computed for a range of $\lambda$. This implicitly changes the effective bandwidths $(\frac{b}{\lambda}, \frac{b}{1-\lambda})$ as $\lambda$ varies. This results in a built-in mechanism where bigger bandwidth ratio is used as the sample size of a regime shrinks. One might want to make a comparison with the test based on $Wald^{(S)}$. But note that one needs to adjust the bandwidth ratios $(b_1, b_2)$ every time $\lambda$ changes so that $(b_1, b_2)$ stay equal to $(\frac{b}{\lambda}, \frac{b}{1-\lambda})$. The implementation of tests with $Wald^{(S)}$ is omitted in this chapter. Denote the $Wald^{(F)}$ statistic computed using break date $T_b \equiv [\lambda T]$ by $Wald^{(F)}(T_b)$. Also denote the limit of $Wald^{(F)}(T_b)$ as $Wald_{\infty}^{(F)}(\lambda)$ where the form of $Wald_{\infty}^{(F)}(\lambda)$ depends on whether traditional or fixed-$b$ asymptotic theory is being used. In the case of fixed-$b$ theory, $Wald_{\infty}^{(F)}(\lambda)$ depends on $\mathbf{P}\left(b, F_p\left(r, \lambda\right)\right)$. As argued by Andrews (1993) and Andrews and Ploberger (1994), break dates close to the end points of the sample cannot be used and so some trimming is needed. To that end define $\Xi^* = [\epsilon T, T - \epsilon T]$ with $0 < \epsilon < 1$ to be the set of admissible break dates. The tuning parameter $\epsilon$ denotes the amount of trimming of potential break dates. Consider the three statistics following Andrews (1993) and Andrews and Ploberger (1994) defined as

$$SupW^{(F)} \equiv \sup_{T_b \in \Xi^*} Wald^{(F)}(T_b), \tag{1.37}$$

$$MeanW^{(F)} \equiv \frac{1}{T} \sum_{T_b \in \Xi^*} Wald^{(F)}(T_b), \tag{1.38}$$

$$ExpW^{(F)} \equiv \log\left(\frac{1}{T}\sum_{T_b\in\Xi^*}\exp\left[\frac{1}{2}Wald^{(F)}(T_b)\right]\right). \tag{1.39}$$

The asymptotic limits of these statistics follow from the continuous mapping theorem and are given by

$$\begin{aligned}
SupW^{(F)} &\Rightarrow \sup_{\lambda\in(\epsilon,1-\epsilon)} Wald_{\infty}^{(F)}(\lambda),\\
MeanW^{(F)} &\Rightarrow \int_{\epsilon}^{1-\epsilon} Wald_{\infty}^{(F)}(\lambda)d\lambda,\\
ExpW^{(F)} &\Rightarrow \log\left(\int_{\epsilon}^{1-\epsilon}\exp\left[\frac{1}{2}Wald_{\infty}^{(F)}(\lambda)\right]d\lambda\right).
\end{aligned}$$

In Tables 1.26 and 1.27 fixed-$b$ critical values for $SupW^{(F)}$, $MeanW^{(F)}$, and $ExpW^{(F)}$ are provided for $l = 2$, $\epsilon = 0.05, 0.1, 0.2$ and for $b \in \{0.02, 0.04, 0.06, 0.08, 0.1, 0.2, 0.3, ..., 0.9, 1\}$.

Tables 1.28 through 1.36 presents the simulation result for some of the DGP specifications introduced in Section 1.6 and for T=100, 500 and 1000. Fixed-$b$ critical values are used for the Bartlett and QS kernels. For the Bartlett kernel, results are also reported using the traditional critical values obtained by Andrews (1993) and Andrews and Ploberger (1994). Several patterns stand out for the null rejection probabilities associated with $SupW^{(F)}$ in Table 1.28 to 1.30. First, rejections using the traditional critical values are often substantially above the 5% nominal level unless persistence is very weak and a small bandwidth is used. Rejections can be close to 100%. The situation is much improved by the use of fixed-$b$ critical values but severe over-rejections are still possible. Size distortions are higher with more persistence in the data, with a smaller value of $\epsilon$, and a smaller value of $b$. As was true in the case of a known break date, the QS kernel gives less size distortion than the Bartlett kernel although the use of large bandwidths causes under-rejections. But over-rejections and under-rejections dissipate as $T$ grows. The Bartlett kernel can suffer from over-rejections that are not easily removed just by using a big bandwidth. A larger value of $T$ helps along with more trimming. Similar patterns hold for the $MeanW^{(F)}$ and $ExpW^{(F)}$ statistics; see Tables 1.31-1.33 and 1.34-1.36 respectively.

## 1.10 Summary and Conclusions

In this chapter fixed-$b$ asymptotics was applied to the problem of testing for the presence of a structural break in a weakly dependent time series regression. Two different HAC estimators and accordingly two different Wald statistics were investigated. The $Wald^{(F)}$ statistic is the Wald statistic that one obtains when structural change is expressed in terms of dummy variables interacted with regressors. The $Wald^{(S)}$ statistic is a restricted version of $Wald^{(F)}$ where the off-diagonal blocks of the HAC estimator are set to zero mimicking the asymptotic zero covariance between OLS estimators in the two regimes. The fixed-$b$ limits of the two statistics were derived, and the fixed-$b$ inference was compared with the traditional inference. In a model with full structural change, both Wald statistics have pivotal fixed-$b$ limits. However, in models with partial structural change, the straightforwardly adapted version of $Wald^{(F)}$ has the same pivotal fixed-$b$ limit as in the full structural change case whereas the straightforwardly adapted version of $Wald^{(S)}$ does not have pivotal fixed-$b$ limit.

In simulation study the finite sample size distortions associated with the fixed-$b$ approach and the traditional approach were examined. The simulation results indicate that the traditional inference is more subject to severe size distortions. When small bandwidths are used, the gap is not huge but as $b$ gets bigger, the difference becomes substantial. Overall, rejections obtained when using fixed-$b$ critical values are closer to the nominal level compared to using the traditional chi-square critical values. When fixed-$b$ critical values are used, finite sample size distortions becomes more pronounce as $b$ gets smaller or as $\lambda$ gets closer to 0 or 1. Local asymptotic power is decreasing in $b$ and power is highest for structural change located near the center of the sample.

In a comparison of the $Wald^{(F)}$ and $Wald^{(S)}$ statistics for full structural change models it was found that $Wald^{(F)}$ tends to be less size distorted than $Wald^{(S)}$ when using fixed-$b$ critical values especially when serial correlation is strong and a large bandwidth is used. The better size performance of $Wald^{(F)}$ comes at the cost of lower power reflecting the usual trade-off between size robustness and power typically found in fixed-$b$ analyses. The choice between $Wald^{(F)}$ and $Wald^{(S)}$ becomes a choice between tolerance for over-rejections relative to desire for high power. At a practical level, $Wald^{(F)}$ is appealing because it retains the same asymptotic pivotal fixed-$b$ limit in models with partial structural change whereas $Wald^{(S)}$ becomes nonpivotal.

Finally, some fixed-$b$ critical values tabulated for $SupW$, $MeanW$, and $ExpW$ statistics which are commonly used for testing the presence of a structural break when the break date is not known a priori. A simulation study revealed that over-rejections are a bigger concern when the break date is treated as unknown. Critical values based on traditional asymptotics can lead to very severe over-rejection problems. Rejections using fixed-$b$ critical values are less distorted especially when the QS kernel is used with a bandwidth that is not too small. When the Bartlett kernel is used, fixed-$b$ rejections show substantial over-rejections.

Table 1.1: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with Bartlett kernel, $l = 1$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 5.94 | 4.68 | 4.34 | 4.22 | 4.2 | 4.21 | 4.31 | 4.59 | 5.92 |
| 0.04 | 8.36 | 5.59 | 4.85 | 4.58 | 4.52 | 4.56 | 4.78 | 5.44 | 8.28 |
| 0.06 | 10.69 | 6.45 | 5.33 | 4.97 | 4.89 | 4.95 | 5.32 | 6.26 | 10.55 |
| 0.08 | 12.68 | 7.37 | 5.94 | 5.39 | 5.21 | 5.33 | 5.96 | 7.21 | 12.67 |
| 0.1 | 14.77 | 8.26 | 6.55 | 5.86 | 5.61 | 5.78 | 6.52 | 8.21 | 14.81 |
| 0.2 | 23.82 | 13 | 9.75 | 8.49 | 8.06 | 8.37 | 9.74 | 12.89 | 24.15 |
| 0.3 | 32.48 | 17.66 | 13.6 | 11.79 | 11.19 | 11.78 | 13.64 | 17.73 | 32.85 |
| 0.4 | 41.16 | 22.74 | 17.5 | 15.22 | 14.59 | 15.18 | 17.62 | 22.72 | 41.63 |
| 0.5 | 50.17 | 27.59 | 21.61 | 18.81 | 18.85 | 18.75 | 21.77 | 27.98 | 50.61 |
| 0.6 | 58.98 | 32.89 | 25.87 | 22.28 | 21.89 | 22.39 | 25.86 | 32.9 | 59.99 |
| 0.7 | 68.93 | 37.77 | 29.8 | 25.79 | 25.3 | 25.97 | 29.98 | 38.06 | 69.38 |
| 0.8 | 77.9 | 42.82 | 33.54 | 29.52 | 28.36 | 29.37 | 34.03 | 43.43 | 78.76 |
| 0.9 | 87.48 | 47.7 | 37.35 | 32.84 | 31.88 | 32.59 | 37.78 | 48.04 | 88.14 |
| 1 | 97.25 | 53.15 | 41.77 | 36.3 | 35.58 | 36.3 | 41.77 | 53.15 | 97.25 |

Table 1.2: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with Bartlett kernel, $l = 2$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 10.45 | 7.49 | 6.74 | 6.49 | 6.53 | 6.6 | 7.04 | 7.78 | 10.73 |
| 0.04 | 15.43 | 9.36 | 7.74 | 7.2 | 7.18 | 7.34 | 8.1 | 9.74 | 15.9 |
| 0.06 | 20.01 | 11.3 | 8.86 | 7.98 | 7.93 | 8.14 | 9.21 | 11.64 | 20.46 |
| 0.08 | 23.63 | 13.23 | 10.08 | 8.92 | 8.76 | 9 | 10.41 | 13.68 | 24.46 |
| 0.1 | 27.45 | 15.17 | 11.16 | 9.77 | 9.62 | 9.97 | 11.86 | 15.61 | 28.41 |
| 0.2 | 45.04 | 24.5 | 18.28 | 15.5 | 15.22 | 15.8 | 18.77 | 25.29 | 45.98 |
| 0.3 | 62.32 | 34.33 | 25.71 | 22.44 | 21.98 | 22.62 | 26.46 | 35.39 | 63.27 |
| 0.4 | 79.15 | 44.31 | 34.01 | 29.69 | 28.71 | 29.56 | 34.46 | 46.08 | 81.68 |
| 0.5 | 97.5 | 54.75 | 41.9 | 37.15 | 36 | 36.28 | 42.98 | 56.62 | 100.45 |
| 0.6 | 115.81 | 64.44 | 50.3 | 44.32 | 42.62 | 44.31 | 51.09 | 67.64 | 119.13 |
| 0.7 | 134.73 | 75.73 | 58.44 | 50.96 | 49.05 | 51.09 | 59.28 | 78.46 | 137.84 |
| 0.8 | 153.2 | 85.47 | 65.2 | 56.85 | 55.35 | 56.79 | 66.47 | 88.5 | 155.94 |
| 0.9 | 171.39 | 94.64 | 71.96 | 63.23 | 61.05 | 63.14 | 73.97 | 98.19 | 174.38 |
| 1 | 197.92 | 111.33 | 84.82 | 74.22 | 71.25 | 74.22 | 84.82 | 111.33 | 197.92 |

Table 1.3: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with Parzen kernel, $l = 1$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 5.53 | 4.48 | 4.22 | 4.12 | 4.13 | 4.14 | 4.2 | 4.42 | 5.45 |
| 0.04 | 7.74 | 5.21 | 4.63 | 4.43 | 4.36 | 4.41 | 4.61 | 5.11 | 7.62 |
| 0.06 | 10.05 | 5.99 | 5.08 | 4.71 | 4.66 | 4.7 | 5 | 5.83 | 10.05 |
| 0.08 | 12.8 | 6.78 | 5.48 | 5.04 | 4.93 | 5.01 | 5.46 | 6.59 | 12.6 |
| 0.1 | 15.39 | 7.61 | 5.93 | 5.39 | 5.19 | 5.34 | 5.97 | 7.44 | 15.31 |
| 0.2 | 28.72 | 12.88 | 8.9 | 7.53 | 7.13 | 7.47 | 8.82 | 12.53 | 29.14 |
| 0.3 | 43.41 | 18.92 | 12.78 | 10.38 | 9.81 | 10.52 | 12.78 | 18.77 | 44.02 |
| 0.4 | 59.19 | 26.32 | 17.63 | 14.4 | 13.6 | 14.79 | 17.71 | 26.38 | 60.09 |
| 0.5 | 77.93 | 35.19 | 24.52 | 19.48 | 18.73 | 19.75 | 24.37 | 35.61 | 78.96 |
| 0.6 | 101.1 | 45.41 | 32.72 | 25.83 | 25.14 | 26.31 | 32.14 | 46.83 | 101.31 |
| 0.7 | 126.46 | 58.14 | 41.87 | 33.32 | 32.72 | 34.24 | 41.36 | 58.84 | 126.34 |
| 0.8 | 156.18 | 72.5 | 53.07 | 42.13 | 42.21 | 43.55 | 52.62 | 74.23 | 154.86 |
| 0.9 | 193.5 | 88.33 | 65.16 | 52.59 | 52.45 | 53.96 | 64.97 | 91.91 | 189.83 |
| 1 | 236.76 | 106.98 | 78.8 | 64.15 | 65.05 | 65.94 | 79.76 | 112.21 | 229.3 |

Table 1.4: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with Parzen kernel, $l = 2$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 9.68 | 7.16 | 6.54 | 6.32 | 6.43 | 6.44 | 6.81 | 7.48 | 9.87 |
| 0.04 | 14.7 | 8.68 | 7.33 | 6.9 | 6.92 | 7.01 | 7.68 | 9.04 | 15.26 |
| 0.06 | 20.78 | 10.44 | 8.22 | 7.5 | 7.46 | 7.62 | 8.6 | 10.85 | 21.32 |
| 0.08 | 26.9 | 12.28 | 9.19 | 8.17 | 8.13 | 8.37 | 9.59 | 12.71 | 27.69 |
| 0.1 | 32.86 | 14.32 | 10.3 | 8.96 | 8.73 | 9.13 | 10.68 | 14.79 | 33.72 |
| 0.2 | 61.88 | 26.88 | 17.11 | 13.96 | 13.6 | 14.16 | 17.92 | 27.25 | 62.55 |
| 0.3 | 95.45 | 42.82 | 27.94 | 22.48 | 21.23 | 22.38 | 28.87 | 44.02 | 97.73 |
| 0.4 | 139.66 | 65.29 | 42.75 | 35.2 | 32.89 | 34.8 | 43.65 | 66.36 | 144.3 |
| 0.5 | 196.8 | 95.6 | 63.74 | 53.21 | 49.54 | 51.62 | 64.56 | 97.41 | 201.36 |
| 0.6 | 278.36 | 134.62 | 92.64 | 77.72 | 72.15 | 75.69 | 92.71 | 137.41 | 282.96 |
| 0.7 | 382.08 | 186.56 | 131.83 | 110.23 | 101.57 | 105.94 | 128.55 | 190.44 | 385.63 |
| 0.8 | 514.04 | 248.31 | 179.83 | 150.7 | 139.6 | 145.76 | 173.65 | 259.83 | 520.59 |
| 0.9 | 680.53 | 326.98 | 234.32 | 200.3 | 184.72 | 194.26 | 230.37 | 341.18 | 702.39 |
| 1 | 892.26 | 423.95 | 303.13 | 260.04 | 240.38 | 251.93 | 303.92 | 435.32 | 914.14 |

Table 1.5: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with QS kernel, $l = 1$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 7.43 | 5.07 | 4.52 | 4.36 | 4.29 | 4.34 | 4.53 | 4.98 | 7.31 |
| 0.04 | 12.21 | 6.48 | 5.28 | 4.91 | 4.79 | 4.86 | 5.27 | 6.34 | 12.04 |
| 0.06 | 17.56 | 8.08 | 6.17 | 5.54 | 5.3 | 5.46 | 6.22 | 7.97 | 17.35 |
| 0.08 | 22.64 | 9.95 | 7.23 | 6.23 | 5.99 | 6.23 | 7.24 | 9.82 | 22.78 |
| 0.1 | 27.85 | 12.16 | 8.43 | 7.08 | 6.76 | 6.99 | 8.43 | 11.95 | 28.47 |
| 0.2 | 57.72 | 25.61 | 16.86 | 13.44 | 12.75 | 13.88 | 16.89 | 25.34 | 59.62 |
| 0.3 | 99.95 | 45.22 | 32.2 | 25.09 | 24.18 | 25.58 | 31.73 | 46.97 | 101.36 |
| 0.4 | 158.39 | 74.89 | 54.62 | 43.04 | 43.29 | 44.15 | 53.58 | 76.5 | 158.68 |
| 0.5 | 249.48 | 114.08 | 83.55 | 68.85 | 70.81 | 70.42 | 85.95 | 118.39 | 239.27 |
| 0.6 | 367.72 | 168.98 | 122.86 | 103.11 | 103.31 | 103.7 | 128.44 | 175.59 | 353.19 |
| 0.7 | 532.6 | 241.2 | 176.66 | 143.09 | 146.09 | 148.06 | 182.71 | 247.64 | 501.66 |
| 0.8 | 726.47 | 326.56 | 245.78 | 197.2 | 199.56 | 201.11 | 247.71 | 343.69 | 697.25 |
| 0.9 | 970.42 | 431.92 | 323.43 | 262.52 | 262.71 | 262.84 | 325.08 | 452.05 | 926.69 |
| 1 | 1261.34 | 559.48 | 416.24 | 339.54 | 340.36 | 334.59 | 423.93 | 585.82 | 1216.77 |

Table 1.6: 95% Fixed-$b$ critical values of $Wald^{(F)}$ with QS kernel, $l = 2$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 14.06 | 8.45 | 7.16 | 6.81 | 6.82 | 6.89 | 7.56 | 8.73 | 14.73 |
| 0.04 | 26.97 | 11.83 | 8.83 | 7.95 | 7.88 | 8.08 | 9.28 | 12.26 | 27.58 |
| 0.06 | 38.59 | 15.95 | 11.06 | 9.39 | 9.18 | 9.58 | 11.44 | 16.41 | 39.04 |
| 0.08 | 50.18 | 20.7 | 13.43 | 11.1 | 10.88 | 11.25 | 14.14 | 21.28 | 50.94 |
| 0.1 | 62.26 | 25.97 | 16.55 | 13.27 | 12.82 | 13.42 | 17.24 | 26.42 | 62.99 |
| 0.2 | 146.13 | 67.5 | 43.02 | 35.24 | 33.63 | 35.2 | 44.36 | 69.12 | 149.5 |
| 0.3 | 312.09 | 158.42 | 108.51 | 87.98 | 80.79 | 87.85 | 106.46 | 163.47 | 314.22 |
| 0.4 | 658.6 | 331.17 | 241.54 | 196.85 | 184.84 | 195.45 | 233.84 | 339.63 | 665.6 |
| 0.5 | 1286.25 | 655.18 | 499.07 | 420.69 | 378.64 | 404.62 | 478.49 | 685.69 | 1357.87 |
| 0.6 | 2448.94 | 1237.59 | 916.23 | 798.85 | 722.79 | 773.38 | 897.04 | 1264.11 | 2526.8 |
| 0.7 | 4375.15 | 2209.76 | 1622.99 | 1431.74 | 1296.51 | 1368.93 | 1595.1 | 2240.27 | 4503.02 |
| 0.8 | 7288.27 | 3760.93 | 2728.88 | 2359.81 | 2208.22 | 2208.18 | 2662.06 | 3808.74 | 7606.67 |
| 0.9 | 11611.96 | 5871.16 | 4433.2 | 3809.97 | 3646.67 | 3495.78 | 4198.07 | 6017.23 | 12269.85 |
| 1 | 17823.91 | 8876.78 | 6763.38 | 5763.21 | 5547.04 | 5223.44 | 6474.23 | 9157.09 | 18845.07 |

Table 1.7: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with Bartlett kernel, $l = 1$, $b_1 = b_2 = b$

| | $\lambda$ = 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$ = 0.02 | 4.101 | 4.025 | 4.027 | 4.022 | 4.033 | 4.046 | 3.991 | 4.019 | 4.115 |
| 0.04 | 4.313 | 4.207 | 4.174 | 4.183 | 4.202 | 4.195 | 4.147 | 4.189 | 4.267 |
| 0.06 | 4.544 | 4.38 | 4.365 | 4.323 | 4.334 | 4.363 | 4.327 | 4.326 | 4.46 |
| 0.08 | 4.71 | 4.569 | 4.543 | 4.503 | 4.517 | 4.513 | 4.507 | 4.519 | 4.621 |
| 0.1 | 4.933 | 4.773 | 4.757 | 4.668 | 4.691 | 4.702 | 4.714 | 4.697 | 4.862 |
| 0.2 | 6.14 | 5.925 | 5.723 | 5.682 | 5.571 | 5.6 | 5.67 | 5.768 | 6.107 |
| 0.3 | 7.65 | 7.112 | 6.991 | 6.727 | 6.728 | 6.566 | 6.93 | 6.965 | 7.51 |
| 0.4 | 9.16 | 8.422 | 8.281 | 8.086 | 7.931 | 7.835 | 8.198 | 8.339 | 9.141 |
| 0.5 | 10.88 | 9.928 | 9.666 | 9.453 | 9.387 | 9.317 | 9.697 | 9.823 | 10.774 |
| 0.6 | 12.62 | 11.477 | 11.063 | 10.91 | 10.886 | 10.879 | 11.258 | 11.472 | 12.538 |
| 0.7 | 14.39 | 13.215 | 12.677 | 12.493 | 12.541 | 12.409 | 12.916 | 13.08 | 14.292 |
| 0.8 | 16.23 | 14.886 | 14.418 | 14.16 | 14.231 | 14.103 | 14.562 | 14.771 | 16.041 |
| 0.9 | 18.28 | 16.586 | 16.083 | 15.794 | 16.021 | 15.764 | 16.414 | 16.613 | 17.992 |
| 1 | 20.31 | 18.425 | 17.886 | 17.582 | 17.792 | 17.521 | 18.177 | 18.41 | 20.008 |

Table 1.8: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with Bartlett kernel, $l = 2$, $b_1 = b_2 = b$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 6.416 | 6.199 | 6.196 | 6.122 | 6.246 | 6.197 | 6.428 | 6.481 | 6.557 |
| 0.04 | 6.788 | 6.557 | 6.459 | 6.404 | 6.527 | 6.553 | 6.702 | 6.835 | 6.987 |
| 0.06 | 7.241 | 6.916 | 6.786 | 6.739 | 6.851 | 6.867 | 7.06 | 7.22 | 7.422 |
| 0.08 | 7.695 | 7.34 | 7.114 | 7.061 | 7.183 | 7.184 | 7.471 | 7.588 | 7.865 |
| 0.1 | 8.224 | 7.774 | 7.461 | 7.395 | 7.518 | 7.561 | 7.845 | 8.024 | 8.388 |
| 0.2 | 10.957 | 10.138 | 9.728 | 9.466 | 9.526 | 9.587 | 10.041 | 10.526 | 11.256 |
| 0.3 | 14.159 | 12.953 | 12.297 | 11.799 | 11.982 | 12.059 | 12.699 | 13.346 | 14.675 |
| 0.4 | 17.909 | 16.119 | 15.067 | 14.606 | 14.84 | 14.835 | 15.726 | 16.787 | 18.563 |
| 0.5 | 21.74 | 19.6 | 18.337 | 17.7 | 17.746 | 17.997 | 18.904 | 20.126 | 22.678 |
| 0.6 | 25.701 | 22.974 | 21.6 | 20.891 | 20.907 | 20.952 | 22.073 | 23.75 | 26.415 |
| 0.7 | 29.558 | 26.374 | 24.714 | 23.974 | 24.045 | 23.829 | 25.325 | 27.618 | 30.453 |
| 0.8 | 33.527 | 29.887 | 28.008 | 27.221 | 27.304 | 26.967 | 28.709 | 30.806 | 34.488 |
| 0.9 | 37.692 | 33.535 | 31.375 | 30.147 | 30.581 | 30.212 | 32.187 | 34.566 | 38.466 |
| 1 | 41.749 | 37.265 | 34.985 | 33.507 | 33.784 | 33.558 | 35.802 | 38.393 | 42.774 |

Table 1.9: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with Parzen kernel, $l = 1$, $b_1 = b_2 = b$

| | $\lambda$= 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 4.049 | 3.982 | 3.999 | 3.986 | 4.002 | 4.003 | 3.963 | 3.978 | 4.047 |
| 0.04 | 4.218 | 4.121 | 4.124 | 4.106 | 4.136 | 4.111 | 4.096 | 4.1 | 4.185 |
| 0.06 | 4.378 | 4.258 | 4.258 | 4.236 | 4.251 | 4.245 | 4.194 | 4.245 | 4.347 |
| 0.08 | 4.549 | 4.407 | 4.38 | 4.356 | 4.353 | 4.371 | 4.349 | 4.346 | 4.484 |
| 0.1 | 4.709 | 4.559 | 4.532 | 4.46 | 4.509 | 4.492 | 4.496 | 4.503 | 4.646 |
| 0.2 | 5.673 | 5.439 | 5.326 | 5.242 | 5.188 | 5.187 | 5.27 | 5.307 | 5.582 |
| 0.3 | 6.89 | 6.457 | 6.209 | 6.132 | 6.057 | 6.013 | 6.2 | 6.282 | 6.811 |
| 0.4 | 8.337 | 7.569 | 7.332 | 7.127 | 7.094 | 6.996 | 7.323 | 7.478 | 8.278 |
| 0.5 | 9.952 | 8.899 | 8.604 | 8.408 | 8.21 | 8.098 | 8.571 | 8.742 | 9.905 |
| 0.6 | 11.766 | 10.479 | 10.063 | 9.749 | 9.586 | 9.516 | 10.086 | 10.371 | 11.863 |
| 0.7 | 14.238 | 12.371 | 11.714 | 11.269 | 11.267 | 11.291 | 11.839 | 12.212 | 14.041 |
| 0.8 | 16.803 | 14.548 | 13.621 | 13.113 | 13.106 | 13.188 | 13.911 | 14.335 | 16.533 |
| 0.9 | 19.604 | 17.11 | 15.697 | 15.141 | 15.394 | 15.483 | 16.075 | 16.557 | 19.562 |
| 1 | 22.963 | 19.776 | 18.188 | 17.469 | 17.794 | 17.65 | 18.511 | 19.084 | 22.964 |

Table 1.10: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with Parzen kernel, $l = 2$, $b_1 = b_2 = b$

| | $\lambda$=0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 6.318 | 6.118 | 6.118 | 6.064 | 6.157 | 6.145 | 6.333 | 6.423 | 6.435 |
| 0.04 | 6.593 | 6.427 | 6,368 | 6.266 | 6.425 | 6.377 | 6.587 | 6.688 | 6.782 |
| 0.06 | 6.926 | 6.723 | 6.558 | 6.519 | 6.658 | 6.658 | 6.841 | 6.992 | 7.14 |
| 0.08 | 7.314 | 7.027 | 6.83 | 6.796 | 6.918 | 6.911 | 7.14 | 7.269 | 7.52 |
| 0.1 | 7.734 | 7.344 | 7.109 | 7.08 | 7.194 | 7.169 | 7.438 | 7.597 | 7.923 |
| 0.2 | 9.978 | 9.293 | 8.823 | 8.604 | 8.775 | 8.794 | 9.174 | 9.51 | 10.252 |
| 0.3 | 13.036 | 11.711 | 11.071 | 10.663 | 10.725 | 10.809 | 11.466 | 12.068 | 13.438 |
| 0.4 | 16.699 | 14.885 | 13.781 | 13.07 | 13.217 | 13.365 | 14.299 | 15.271 | 17.304 |
| 0.5 | 21.601 | 18.733 | 17.057 | 16.273 | 16.441 | 16.498 | 17.645 | 19.301 | 22.368 |
| 0.6 | 27.984 | 23.528 | 21.279 | 20.453 | 20.275 | 20.393 | 22.011 | 24.447 | 28.489 |
| 0.7 | 35.876 | 29.316 | 26.426 | 25.412 | 25.186 | 25.205 | 27.459 | 30.67 | 36.487 |
| 0.8 | 45.449 | 36.522 | 32.699 | 31.458 | 30.822 | 30.985 | 34.199 | 38.416 | 45.707 |
| 0.9 | 56.114 | 44.869 | 40.393 | 38.288 | 37.563 | 37.654 | 41.742 | 47.109 | 57.801 |
| 1 | 68.506 | 54.481 | 48.77 | 45.915 | 45.317 | 45.607 | 49.919 | 57.401 | 70.667 |

Table 1.11: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with QS kernel, $l = 1$, $b_1 = b_2 = b$

| | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 4.176 | 4.097 | 4.082 | 4.082 | 4.098 | 4.078 | 4.065 | 4.073 | 4.162 |
| 0.04 | 4.466 | 4.324 | 4.33 | 4.303 | 4.299 | 4.298 | 4.281 | 4.301 | 4.418 |
| 0.06 | 4.789 | 4.629 | 4.572 | 4.515 | 4.542 | 4.534 | 4.537 | 4.551 | 4.709 |
| 0.08 | 5.109 | 4.91 | 4.858 | 4.756 | 4.792 | 4.776 | 4.816 | 4.817 | 5.027 |
| 0.1 | 5.461 | 5.239 | 5.15 | 5.055 | 5.045 | 5.031 | 5.095 | 5.121 | 5.376 |
| 0.2 | 7.908 | 7.158 | 6.91 | 6.746 | 6.723 | 6.603 | 6.924 | 7.033 | 7.796 |
| 0.3 | 11.14 | 9.864 | 9.397 | 9.007 | 8.873 | 8.798 | 9.357 | 9.723 | 11.1 |
| 0.4 | 15.772 | 13.486 | 12.522 | 12.13 | 12.109 | 12.24 | 12.867 | 13.339 | 15.716 |
| 0.5 | 21.421 | 18.5 | 16.806 | 16.16 | 16.533 | 16.272 | 17.009 | 17.961 | 21.766 |
| 0.6 | 28.285 | 24.521 | 22.728 | 21.041 | 21.838 | 21.285 | 22.432 | 24.054 | 29.2 |
| 0.7 | 37.351 | 32.408 | 29.597 | 27.6 | 28.351 | 27.798 | 29.577 | 32.03 | 38.786 |
| 0.8 | 48.093 | 41.538 | 38.324 | 35.403 | 36.689 | 35.926 | 38.133 | 41.491 | 49.723 |
| 0.9 | 60.429 | 52.591 | 48.937 | 44.754 | 46.413 | 46.106 | 48.34 | 52.617 | 62.802 |
| 1 | 75.711 | 64.723 | 60.735 | 55.78 | 57.464 | 56.974 | 60.451 | 66.712 | 78.41 |

Table 1.12: 95% Fixed-$b$ critical values of $Wald^{(S)}$ with QS kernel, $l = 2$, $b_1 = b_2 = b$

| | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|---|---|---|---|---|
| $b$=0.02 | 6.545 | 6.371 | 6.319 | 6.233 | 6.359 | 6.326 | 6.53 | 6.635 | 6.69 |
| 0.04 | 7.194 | 6.916 | 6.741 | 6.662 | 6.792 | 6.83 | 7.014 | 7.19 | 7.38 |
| 0.06 | 7.939 | 7.533 | 7.247 | 7.182 | 7.315 | 7.316 | 7.55 | 7.747 | 8.162 |
| 0.08 | 8.715 | 8.208 | 7.811 | 7.69 | 7.836 | 7.875 | 8.242 | 8.427 | 8.986 |
| 0.1 | 9.665 | 8.968 | 8.486 | 8.306 | 8.462 | 8.466 | 8.922 | 9.2 | 9.933 |
| 0.2 | 16.041 | 13.952 | 13.07 | 12.468 | 12.481 | 12.526 | 13.544 | 14.544 | 16.558 |
| 0.3 | 27.467 | 22.471 | 20.156 | 19.41 | 19.266 | 19.249 | 21.214 | 23.57 | 27.843 |
| 0.4 | 45.53 | 35.924 | 31.708 | 30.315 | 29.839 | 30.121 | 33.261 | 37.764 | 46.608 |
| 0.5 | 72.151 | 54.831 | 49.151 | 45.982 | 45.157 | 45.123 | 50.044 | 57.892 | 74.435 |
| 0.6 | 108.625 | 82.168 | 72.852 | 66.928 | 66.757 | 66.24 | 72.815 | 86.59 | 112.953 |
| 0.7 | 158.481 | 117.545 | 103.562 | 96.266 | 94.481 | 93.173 | 104.123 | 125.598 | 162.34 |
| 0.8 | 221.964 | 163.498 | 143.644 | 133.73 | 129.628 | 129.225 | 144.891 | 175.599 | 227.359 |
| 0.9 | 304.042 | 220.842 | 193.648 | 182.573 | 174.469 | 174.765 | 193.792 | 242.272 | 308.793 |
| 1 | 404.653 | 296.357 | 256.175 | 243.928 | 231.015 | 229.871 | 256.899 | 316.291 | 412.891 |

Table 1.13: 95% Fixed-$b$ critical values of $Wald^{(S)}$, $l = 2$, $b_1 \neq b_2$

| $\lambda$ | $b_1$ | $b_2$ | Bartlett kernel | QS kernel |
|---|---|---|---|---|
| 0.2 | 0.2 | 0.05 | 9.2936 | 11.8473 |
| 0.2 | 0.5 | 0.125 | 14.9354 | 24.8345 |
| 0.2 | 1 | 0.25 | 23.8607 | 59.9221 |
| 0.4 | 1.25 | 0.83 | 34.864 | 256.3 |
| 0.5 | 1 | 1 | 33.7844 | 231.0145 |
| 0.5 | 2 | 2 | 67.5688 | 1677.88 |

Table 1.14: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP A: $(\theta, \rho, \varphi) = (0.5, 0, 0)$.
$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.2$

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-$b$ | 0.1784 | 0.1636 | 0.1544 | 0.1444 | 0.1348 | 0.1256 | 0.122 |
| | Bartlett | chi square | 0.1892 | 0.1916 | 0.1948 | 0.1984 | 0.2008 | 0.2544 | 0.3088 |
| | QS | Fixed-$b$ | 0.1704 | 0.1548 | 0.1372 | 0.1236 | 0.1124 | 0.1064 | 0.0964 |
| | QS | chi square | 0.1888 | 0.194 | 0.1936 | 0.2012 | 0.2132 | 0.2988 | 0.3936 |
| 100 | Bartlett | Fixed-$b$ | 0.1096 | 0.1024 | 0.098 | 0.1 | 0.098 | 0.0928 | 0.0888 |
| | Bartlett | chi square | 0.1176 | 0.1216 | 0.1296 | 0.1428 | 0.1544 | 0.2096 | 0.2692 |
| | QS | Fixed-$b$ | 0.106 | 0.0936 | 0.0916 | 0.0884 | 0.0896 | 0.0892 | 0.0804 |
| | QS | chi square | 0.1216 | 0.1204 | 0.1356 | 0.1596 | 0.1764 | 0.274 | 0.3688 |
| 500 | Bartlett | Fixed-$b$ | 0.066 | 0.064 | 0.0636 | 0.0632 | 0.064 | 0.0632 | 0.0696 |
| | Bartlett | chi square | 0.0724 | 0.0832 | 0.0924 | 0.1036 | 0.1104 | 0.1664 | 0.23 |
| | QS | Fixed-$b$ | 0.0628 | 0.064 | 0.0628 | 0.062 | 0.062 | 0.0668 | 0.062 |
| | QS | chi square | 0.0756 | 0.0916 | 0.1052 | 0.1224 | 0.142 | 0.2284 | 0.3304 |

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-$b$ | 0.1212 | 0.114 | 0.1144 | 0.1176 | 0.1168 | 0.116 | 0.116 |
| | Bartlett | chi square | 0.364 | 0.4152 | 0.464 | 0.5052 | 0.5416 | 0.5724 | 0.5932 |
| | QS | Fixed-$b$ | 0.0916 | 0.0892 | 0.0856 | 0.0868 | 0.0864 | 0.0856 | 0.0848 |
| | QS | chi square | 0.4932 | 0.5764 | 0.6404 | 0.7016 | 0.742 | 0.7776 | 0.8084 |
| 100 | Bartlett | Fixed-$b$ | 0.0876 | 0.0836 | 0.0832 | 0.0876 | 0.0864 | 0.0868 | 0.086 |
| | Bartlett | chi square | 0.322 | 0.3644 | 0.4116 | 0.4508 | 0.4964 | 0.5296 | 0.5596 |
| | QS | Fixed-$b$ | 0.072 | 0.072 | 0.0636 | 0.0644 | 0.0672 | 0.0672 | 0.0656 |
| | QS | chi square | 0.4636 | 0.5408 | 0.6176 | 0.6808 | 0.726 | 0.7676 | 0.7984 |
| 500 | Bartlett | Fixed-$b$ | 0.0696 | 0.0668 | 0.0704 | 0.0672 | 0.0652 | 0.0652 | 0.0652 |
| | Bartlett | chi square | 0.2824 | 0.3316 | 0.3816 | 0.4232 | 0.4664 | 0.5068 | 0.5396 |
| | QS | Fixed-$b$ | 0.0608 | 0.0616 | 0.062 | 0.0628 | 0.062 | 0.0628 | 0.0604 |
| | QS | chi square | 0.4216 | 0.5144 | 0.5964 | 0.6564 | 0.7204 | 0.7692 | 0.8024 |

Table 1.15: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP A: $(\theta, \rho, \varphi) = (0.5, 0, 0)$.

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.4$

| | | | | | | *b* | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-*b* | 0.1036 | 0.0976 | 0.088 | 0.092 | 0.09 | 0.0852 | 0.0844 |
| | Bartlett | chi square | 0.1088 | 0.1096 | 0.116 | 0.1288 | 0.136 | 0.194 | 0.2492 |
| | QS | Fixed-*b* | 0.0988 | 0.0852 | 0.0848 | 0.0864 | 0.088 | 0.084 | 0.0792 |
| | QS | chi square | 0.1088 | 0.1108 | 0.1252 | 0.1376 | 0.1588 | 0.2504 | 0.3452 |
| 100 | Bartlett | Fixed-*b* | 0.0796 | 0.08 | 0.078 | 0.0772 | 0.0748 | 0.076 | 0.0716 |
| | Bartlett | chi square | 0.0816 | 0.092 | 0.0996 | 0.1068 | 0.12 | 0.1688 | 0.2268 |
| | QS | Fixed-*b* | 0.0776 | 0.078 | 0.0772 | 0.0776 | 0.0772 | 0.074 | 0.0696 |
| | QS | chi square | 0.0812 | 0.0972 | 0.1084 | 0.1244 | 0.1432 | 0.2268 | 0.3124 |
| 500 | Bartlett | Fixed-*b* | 0.0672 | 0.0684 | 0.066 | 0.0668 | 0.068 | 0.0644 | 0.0676 |
| | Bartlett | chi square | 0.0692 | 0.0764 | 0.088 | 0.0968 | 0.1044 | 0.1508 | 0.2052 |
| | QS | Fixed-*b* | 0.0656 | 0.0676 | 0.0656 | 0.0644 | 0.0668 | 0.0628 | 0.0572 |
| | QS | chi square | 0.0712 | 0.0868 | 0.1 | 0.1124 | 0.126 | 0.21 | 0.3032 |

| | | | | | | *b* | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-*b* | 0.0832 | 0.0844 | 0.0852 | 0.0828 | 0.0848 | 0.0848 | 0.0852 |
| | Bartlett | chi square | 0.306 | 0.348 | 0.39 | 0.4348 | 0.4708 | 0.506 | 0.5372 |
| | QS | Fixed-*b* | 0.0824 | 0.0808 | 0.0832 | 0.0804 | 0.0764 | 0.076 | 0.0728 |
| | QS | chi square | 0.42 | 0.5048 | 0.5756 | 0.6368 | 0.6884 | 0.7368 | 0.776 |
| 100 | Bartlett | Fixed-*b* | 0.0716 | 0.066 | 0.064 | 0.0656 | 0.066 | 0.0664 | 0.066 |
| | Bartlett | chi square | 0.272 | 0.3268 | 0.374 | 0.4156 | 0.4468 | 0.4764 | 0.5028 |
| | QS | Fixed-*b* | 0.0624 | 0.0604 | 0.0604 | 0.0608 | 0.0616 | 0.0596 | 0.0556 |
| | QS | chi square | 0.3948 | 0.4736 | 0.5372 | 0.598 | 0.656 | 0.7004 | 0.7464 |
| 500 | Bartlett | Fixed-*b* | 0.0588 | 0.062 | 0.0636 | 0.0656 | 0.0664 | 0.0656 | 0.0648 |
| | Bartlett | chi square | 0.2628 | 0.3084 | 0.3512 | 0.39 | 0.432 | 0.4648 | 0.4976 |
| | QS | Fixed-*b* | 0.056 | 0.0644 | 0.064 | 0.0612 | 0.0596 | 0.0588 | 0.0576 |
| | QS | chi square | 0.3828 | 0.4608 | 0.5348 | 0.6028 | 0.658 | 0.7072 | 0.7476 |

Table 1.16: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$.
$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.2$

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-$b$ | 0.4812 | 0.424 | 0.3864 | 0.3548 | 0.336 | 0.2444 | 0.2168 |
| | Bartlett | chi square | 0.4908 | 0.4528 | 0.4348 | 0.4224 | 0.4172 | 0.4068 | 0.4428 |
| | QS | Fixed-$b$ | 0.4712 | 0.3904 | 0.346 | 0.3108 | 0.2756 | 0.178 | 0.1408 |
| | QS | chi square | 0.4928 | 0.4388 | 0.4184 | 0.4092 | 0.3972 | 0.4156 | 0.4812 |
| 100 | Bartlett | Fixed-$b$ | 0.3904 | 0.334 | 0.286 | 0.2468 | 0.2136 | 0.1704 | 0.1504 |
| | Bartlett | chi square | 0.4032 | 0.3648 | 0.334 | 0.3144 | 0.302 | 0.3148 | 0.368 |
| | QS | Fixed-$b$ | 0.3596 | 0.3028 | 0.2332 | 0.192 | 0.1756 | 0.1172 | 0.1044 |
| | QS | chi square | 0.3832 | 0.3516 | 0.3088 | 0.2872 | 0.2808 | 0.3348 | 0.422 |
| 500 | Bartlett | Fixed-$b$ | 0.1892 | 0.124 | 0.106 | 0.0948 | 0.092 | 0.08 | 0.0804 |
| | Bartlett | chi square | 0.1976 | 0.1476 | 0.14 | 0.1428 | 0.1488 | 0.194 | 0.2472 |
| | QS | Fixed-$b$ | 0.1492 | 0.0956 | 0.082 | 0.076 | 0.0716 | 0.0708 | 0.0672 |
| | QS | chi square | 0.164 | 0.1244 | 0.126 | 0.1408 | 0.1544 | 0.2328 | 0.3396 |

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-$b$ | 0.2084 | 0.206 | 0.2036 | 0.2048 | 0.2008 | 0.2008 | 0.2024 |
| | Bartlett | chi square | 0.4832 | 0.5308 | 0.5668 | 0.6088 | 0.6372 | 0.664 | 0.6888 |
| | QS | Fixed-$b$ | 0.1236 | 0.1136 | 0.1056 | 0.1064 | 0.1024 | 0.0996 | 0.096 |
| | QS | chi square | 0.562 | 0.6236 | 0.6888 | 0.7476 | 0.7904 | 0.8248 | 0.8524 |
| 100 | Bartlett | Fixed-$b$ | 0.1464 | 0.1452 | 0.1476 | 0.1464 | 0.144 | 0.1456 | 0.148 |
| | Bartlett | chi square | 0.4184 | 0.47 | 0.5176 | 0.5576 | 0.5944 | 0.6232 | 0.6508 |
| | QS | Fixed-$b$ | 0.0944 | 0.0868 | 0.0844 | 0.084 | 0.082 | 0.082 | 0.0804 |
| | QS | chi square | 0.5104 | 0.5856 | 0.6468 | 0.71 | 0.764 | 0.8048 | 0.8332 |
| 500 | Bartlett | Fixed-$b$ | 0.0832 | 0.08 | 0.0804 | 0.0824 | 0.078 | 0.0804 | 0.0796 |
| | Bartlett | chi square | 0.3048 | 0.36 | 0.4036 | 0.4532 | 0.4956 | 0.5268 | 0.5588 |
| | QS | Fixed-$b$ | 0.0696 | 0.0676 | 0.0668 | 0.0664 | 0.0644 | 0.0636 | 0.0608 |
| | QS | chi square | 0.4332 | 0.5192 | 0.594 | 0.6664 | 0.7236 | 0.7648 | 0.7908 |

Table 1.17: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$.
$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.4$

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-$b$ | 0.4076 | 0.372 | 0.3132 | 0.2564 | 0.2368 | 0.1776 | 0.158 |
| | Bartlett | chi square | 0.416 | 0.3968 | 0.3532 | 0.328 | 0.3092 | 0.3096 | 0.3584 |
| | QS | Fixed-$b$ | 0.4076 | 0.35 | 0.2572 | 0.2184 | 0.1896 | 0.1264 | 0.108 |
| | QS | chi square | 0.422 | 0.3836 | 0.3284 | 0.2956 | 0.284 | 0.3272 | 0.4092 |
| 100 | Bartlett | Fixed-$b$ | 0.3544 | 0.2464 | 0.2036 | 0.1712 | 0.1528 | 0.122 | 0.1164 |
| | Bartlett | chi square | 0.362 | 0.2692 | 0.2328 | 0.2192 | 0.2168 | 0.2364 | 0.2884 |
| | QS | Fixed-$b$ | 0.336 | 0.208 | 0.1528 | 0.1296 | 0.1184 | 0.0904 | 0.0836 |
| | QS | chi square | 0.3492 | 0.24 | 0.2076 | 0.1992 | 0.2012 | 0.2596 | 0.342 |
| 500 | Bartlett | Fixed-$b$ | 0.1116 | 0.0856 | 0.0776 | 0.0764 | 0.0772 | 0.0736 | 0.07 |
| | Bartlett | chi square | 0.1168 | 0.0968 | 0.0956 | 0.1016 | 0.1072 | 0.162 | 0.2124 |
| | QS | Fixed-$b$ | 0.0856 | 0.0688 | 0.068 | 0.0688 | 0.0688 | 0.0636 | 0.0576 |
| | QS | chi square | 0.094 | 0.0888 | 0.0936 | 0.1044 | 0.1204 | 0.202 | 0.2896 |

| | | | $b$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-$b$ | 0.1536 | 0.1512 | 0.1476 | 0.1492 | 0.1512 | 0.1556 | 0.1548 |
| | Bartlett | chi square | 0.4104 | 0.4628 | 0.5112 | 0.5428 | 0.5744 | 0.6032 | 0.6356 |
| | QS | Fixed-$b$ | 0.1004 | 0.1012 | 0.0996 | 0.0992 | 0.0992 | 0.0952 | 0.096 |
| | QS | chi square | 0.4924 | 0.5732 | 0.6324 | 0.6976 | 0.7484 | 0.788 | 0.8152 |
| 100 | Bartlett | Fixed-$b$ | 0.1156 | 0.1072 | 0.1056 | 0.1108 | 0.1124 | 0.112 | 0.1116 |
| | Bartlett | chi square | 0.3372 | 0.3864 | 0.428 | 0.472 | 0.506 | 0.5388 | 0.5744 |
| | QS | Fixed-$b$ | 0.0792 | 0.0784 | 0.0748 | 0.0748 | 0.0728 | 0.0736 | 0.0728 |
| | QS | chi square | 0.422 | 0.5052 | 0.5776 | 0.6392 | 0.6884 | 0.7252 | 0.7696 |
| 500 | Bartlett | Fixed-$b$ | 0.0664 | 0.066 | 0.068 | 0.0668 | 0.0656 | 0.0692 | 0.07 |
| | Bartlett | chi square | 0.2548 | 0.2988 | 0.348 | 0.388 | 0.4308 | 0.4676 | 0.5052 |
| | QS | Fixed-$b$ | 0.0564 | 0.0572 | 0.0576 | 0.0588 | 0.0608 | 0.0592 | 0.0588 |
| | QS | chi square | 0.374 | 0.4588 | 0.5344 | 0.602 | 0.6616 | 0.7016 | 0.7372 |

Table 1.18: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$.
$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.2$

| | | | *b* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-*b* | 0.5204 | 0.462 | 0.4212 | 0.3956 | 0.3728 | 0.2832 | 0.2532 |
| | Bartlett | chi square | 0.5296 | 0.4936 | 0.4712 | 0.46 | 0.4524 | 0.4472 | 0.4836 |
| | QS | Fixed-*b* | 0.5164 | 0.426 | 0.3788 | 0.3428 | 0.3052 | 0.204 | 0.1676 |
| | QS | chi square | 0.5312 | 0.4756 | 0.4548 | 0.4436 | 0.4356 | 0.4608 | 0.5188 |
| 100 | Bartlett | Fixed-*b* | 0.4504 | 0.3984 | 0.3492 | 0.3044 | 0.2728 | 0.2152 | 0.1944 |
| | Bartlett | chi square | 0.4624 | 0.4256 | 0.3948 | 0.3728 | 0.3608 | 0.38 | 0.4168 |
| | QS | Fixed-*b* | 0.4264 | 0.3676 | 0.292 | 0.2464 | 0.2184 | 0.1528 | 0.1292 |
| | QS | chi square | 0.444 | 0.4124 | 0.3688 | 0.3464 | 0.3404 | 0.39 | 0.4588 |
| 500 | Bartlett | Fixed-*b* | 0.2132 | 0.144 | 0.1212 | 0.1084 | 0.1028 | 0.0948 | 0.0916 |
| | Bartlett | chi square | 0.2248 | 0.1676 | 0.1576 | 0.1632 | 0.1692 | 0.218 | 0.28 |
| | QS | Fixed-*b* | 0.1668 | 0.108 | 0.0948 | 0.0872 | 0.0808 | 0.0804 | 0.0804 |
| | QS | chi square | 0.1828 | 0.1428 | 0.1444 | 0.1604 | 0.1772 | 0.2624 | 0.3588 |

| | | | *b* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-*b* | 0.2392 | 0.2328 | 0.2336 | 0.2292 | 0.2284 | 0.2308 | 0.2308 |
| | Bartlett | chi square | 0.5252 | 0.5744 | 0.6128 | 0.6436 | 0.6768 | 0.7028 | 0.726 |
| | QS | Fixed-*b* | 0.1436 | 0.1332 | 0.1268 | 0.1244 | 0.12 | 0.1188 | 0.118 |
| | QS | chi square | 0.5948 | 0.656 | 0.716 | 0.7652 | 0.8076 | 0.8356 | 0.8644 |
| 100 | Bartlett | Fixed-*b* | 0.1884 | 0.1844 | 0.1844 | 0.182 | 0.1828 | 0.1828 | 0.1836 |
| | Bartlett | chi square | 0.4624 | 0.5084 | 0.5572 | 0.596 | 0.6328 | 0.6616 | 0.6908 |
| | QS | Fixed-*b* | 0.1184 | 0.11 | 0.1064 | 0.1016 | 0.102 | 0.1008 | 0.0964 |
| | QS | chi square | 0.5432 | 0.6216 | 0.6868 | 0.744 | 0.7836 | 0.8232 | 0.8488 |
| 500 | Bartlett | Fixed-*b* | 0.0936 | 0.0912 | 0.0916 | 0.0904 | 0.0908 | 0.092 | 0.0928 |
| | Bartlett | chi square | 0.3304 | 0.3836 | 0.4296 | 0.474 | 0.5124 | 0.5476 | 0.5816 |
| | QS | Fixed-*b* | 0.0772 | 0.0768 | 0.0752 | 0.0712 | 0.0692 | 0.0692 | 0.0676 |
| | QS | chi square | 0.4512 | 0.5352 | 0.6156 | 0.6836 | 0.7344 | 0.7744 | 0.8136 |

Table 1.19: The Finite Sample Size associated with $Wald^{(S)}$, $l = 2$, $b_1 = b_2 = b$.

DGP D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$.
$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$, $\lambda = 0.4$

| | | | | | | $b$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 | 0.200 | 0.300 |
| 50 | Bartlett | Fixed-$b$ | 0.4664 | 0.4304 | 0.3632 | 0.3124 | 0.2828 | 0.2144 | 0.2064 |
| | Bartlett | chi square | 0.4768 | 0.4528 | 0.406 | 0.3692 | 0.3512 | 0.3652 | 0.4104 |
| | QS | Fixed-$b$ | 0.4668 | 0.406 | 0.3092 | 0.2536 | 0.2188 | 0.1576 | 0.142 |
| | QS | chi square | 0.4856 | 0.4396 | 0.3708 | 0.3392 | 0.326 | 0.3696 | 0.4496 |
| 100 | Bartlett | Fixed-$b$ | 0.4148 | 0.292 | 0.2368 | 0.2008 | 0.1868 | 0.1528 | 0.1452 |
| | Bartlett | chi square | 0.422 | 0.3128 | 0.2732 | 0.2552 | 0.2492 | 0.2784 | 0.328 |
| | QS | Fixed-$b$ | 0.3964 | 0.2464 | 0.1772 | 0.1532 | 0.1408 | 0.1168 | 0.104 |
| | QS | chi square | 0.4084 | 0.2796 | 0.2416 | 0.2256 | 0.2288 | 0.296 | 0.3792 |
| 500 | Bartlett | Fixed-$b$ | 0.1368 | 0.1 | 0.09 | 0.0896 | 0.0848 | 0.0744 | 0.074 |
| | Bartlett | chi square | 0.142 | 0.1144 | 0.1152 | 0.1228 | 0.1292 | 0.1764 | 0.2296 |
| | QS | Fixed-$b$ | 0.0992 | 0.0812 | 0.0768 | 0.0748 | 0.0736 | 0.0648 | 0.0608 |
| | QS | chi square | 0.1092 | 0.0992 | 0.1088 | 0.1236 | 0.1396 | 0.218 | 0.3068 |

| | | | | | | $b$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| T | kernel | Inference | 0.400 | 0.500 | 0.600 | 0.700 | 0.800 | 0.900 | 1.000 |
| 50 | Bartlett | Fixed-$b$ | 0.1956 | 0.1916 | 0.1872 | 0.1888 | 0.1908 | 0.1916 | 0.1904 |
| | Bartlett | chi square | 0.4584 | 0.5012 | 0.5492 | 0.59 | 0.6244 | 0.6524 | 0.6808 |
| | QS | Fixed-$b$ | 0.132 | 0.1236 | 0.1184 | 0.1112 | 0.1088 | 0.106 | 0.1052 |
| | QS | chi square | 0.5276 | 0.6188 | 0.674 | 0.7256 | 0.7688 | 0.8028 | 0.8328 |
| 100 | Bartlett | Fixed-$b$ | 0.1404 | 0.1336 | 0.1336 | 0.1344 | 0.1344 | 0.1364 | 0.1372 |
| | Bartlett | chi square | 0.376 | 0.4192 | 0.4732 | 0.5148 | 0.5528 | 0.58 | 0.608 |
| | QS | Fixed-$b$ | 0.0968 | 0.0924 | 0.0924 | 0.0912 | 0.0876 | 0.0868 | 0.0864 |
| | QS | chi square | 0.4664 | 0.5576 | 0.624 | 0.6752 | 0.7196 | 0.7588 | 0.7928 |
| 500 | Bartlett | Fixed-$b$ | 0.0708 | 0.0724 | 0.0736 | 0.0744 | 0.074 | 0.0756 | 0.0756 |
| | Bartlett | chi square | 0.2796 | 0.3228 | 0.3732 | 0.4096 | 0.45 | 0.482 | 0.5128 |
| | QS | Fixed-$b$ | 0.0612 | 0.0572 | 0.058 | 0.0604 | 0.0648 | 0.0656 | 0.0648 |
| | QS | chi square | 0.4004 | 0.4768 | 0.5536 | 0.6176 | 0.6716 | 0.7188 | 0.7568 |

Table 1.20: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l = 2$, $b_1 = \frac{b}{\lambda}$, $b_2 = \frac{b}{1-\lambda}$, $T = 50$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .2 | .04 | .2 | .05 | 0.1184 | 0.1168 | 0.2284 | 0.0968 | 0.098 | 0.2664 |
| | | .1 | .5 | .125 | 0.11 | 0.112 | 0.346 | 0.0992 | 0.0928 | 0.464 |
| | | .2 | 1.0 | .25 | 0.1132 | 0.1116 | 0.4924 | 0.0952 | 0.0932 | 0.6596 |
| B | .2 | .04 | .2 | .05 | 0.2444 | 0.244 | 0.3804 | 0.1868 | 0.1884 | 0.3904 |
| | | .1 | .5 | .125 | 0.1704 | 0.17 | 0.436 | 0.1168 | 0.1108 | 0.4936 |
| | | .2 | 1.0 | .25 | 0.1548 | 0.1544 | 0.558 | 0.096 | 0.0932 | 0.6704 |
| C | .2 | .04 | .2 | .05 | 0.2948 | 0.2928 | 0.4244 | 0.2092 | 0.2116 | 0.4152 |
| | | .1 | .5 | .125 | 0.1932 | 0.1952 | 0.4592 | 0.1216 | 0.1184 | 0.506 |
| | | .2 | 1.0 | .25 | 0.1668 | 0.1676 | 0.5728 | 0.0924 | 0.0928 | 0.6668 |
| D | .2 | .04 | .2 | .05 | 0.3296 | 0.3292 | 0.47 | 0.2428 | 0.2444 | 0.4596 |
| | | .1 | .5 | .125 | 0.2312 | 0.2304 | 0.514 | 0.1464 | 0.1416 | 0.5552 |
| | | .2 | 1.0 | .25 | 0.2024 | 0.2036 | 0.6164 | 0.1156 | 0.1168 | 0.6984 |
| E | .2 | .04 | .2 | .05 | 0.5672 | 0.5664 | 0.6812 | 0.4576 | 0.46 | 0.6612 |
| | | .1 | .5 | .125 | 0.3696 | 0.3688 | 0.6444 | 0.2148 | 0.2128 | 0.6412 |
| | | .2 | 1.0 | .25 | 0.2768 | 0.284 | 0.682 | 0.1428 | 0.1384 | 0.712 |
| F | .2 | .04 | .2 | .05 | 0.582 | 0.5804 | 0.6948 | 0.4816 | 0.4872 | 0.6732 |
| | | .1 | .5 | .125 | 0.3924 | 0.3928 | 0.6728 | 0.2492 | 0.2452 | 0.6684 |
| | | .2 | 1.0 | .25 | 0.3064 | 0.3136 | 0.7064 | 0.1696 | 0.1632 | 0.7364 |

Note: The DGP labels are given by A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$, B: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.0)$, C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$, D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$, E: $(\theta, \rho, \varphi) = (0.8, 0.9, 0.5)$, and F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$.

Table 1.21: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l = 2$, $b_1 = \frac{b}{\lambda}$, $b_2 = \frac{b}{1-\lambda}$, $T = 100$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .2 | .04 | .2 | .05 | 0.0928 | 0.092 | 0.1924 | 0.0872 | 0.0892 | 0.2408 |
| | | .1 | .5 | .125 | 0.0876 | 0.0884 | 0.3044 | 0.08 | 0.0796 | 0.43 |
| | | .2 | 1.0 | .25 | 0.0868 | 0.086 | 0.4472 | 0.082 | 0.0784 | 0.6308 |
| B | .2 | .04 | .2 | .05 | 0.1672 | 0.1668 | 0.2916 | 0.1204 | 0.122 | 0.3004 |
| | | .1 | .5 | .125 | 0.1264 | 0.1312 | 0.3872 | 0.088 | 0.086 | 0.4556 |
| | | .2 | 1.0 | .25 | 0.1168 | 0.1152 | 0.5144 | 0.0768 | 0.0716 | 0.6344 |
| C | .2 | .04 | .2 | .05 | 0.1816 | 0.1828 | 0.3132 | 0.1316 | 0.1344 | 0.3124 |
| | | .1 | .5 | .125 | 0.1396 | 0.1376 | 0.4028 | 0.0904 | 0.092 | 0.4624 |
| | | .2 | 1.0 | .25 | 0.1244 | 0.1264 | 0.5288 | 0.0792 | 0.0812 | 0.6424 |
| D | .2 | .04 | .2 | .05 | 0.2312 | 0.2296 | 0.3792 | 0.1636 | 0.168 | 0.3696 |
| | | .1 | .5 | .125 | 0.1772 | 0.1784 | 0.4528 | 0.114 | 0.1108 | 0.5044 |
| | | .2 | 1.0 | .25 | 0.1624 | 0.1636 | 0.58 | 0.1116 | 0.106 | 0.6892 |
| E | .2 | .04 | .2 | .05 | 0.4656 | 0.4672 | 0.6104 | 0.3404 | 0.3428 | 0.5716 |
| | | .1 | .5 | .125 | 0.276 | 0.28 | 0.5848 | 0.1476 | 0.1448 | 0.586 |
| | | .2 | 1.0 | .25 | 0.2192 | 0.23 | 0.6472 | 0.1048 | 0.1032 | 0.6908 |
| F | .2 | .04 | .2 | .05 | 0.4864 | 0.4872 | 0.6308 | 0.3588 | 0.364 | 0.5996 |
| | | .1 | .5 | .125 | 0.3008 | 0.3036 | 0.618 | 0.1764 | 0.1748 | 0.6156 |
| | | .2 | 1.0 | .25 | 0.25 | 0.2524 | 0.6784 | 0.1292 | 0.1232 | 0.7152 |

Note: The DGP labels are given by A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$, B: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.0)$, C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$, D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$, E: $(\theta, \rho, \varphi) = (0.8, 0.9, 0.5)$, and F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$.

Table 1.22: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l=2$, $b_1=\frac{b}{\lambda}$, $b_2=\frac{b}{1-\lambda}$, $T=500$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t=\lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .2 | .04 | .2 | .05 | 0.0676 | 0.0672 | 0.1468 | 0.0628 | 0.0636 | 0.1948 |
| | | .1 | .5 | .125 | 0.066 | 0.0628 | 0.274 | 0.0592 | 0.0616 | 0.3872 |
| | | .2 | 1.0 | .25 | 0.0632 | 0.06 | 0.4128 | 0.0588 | 0.0604 | 0.6104 |
| B | .2 | .04 | .2 | .05 | 0.0788 | 0.078 | 0.174 | 0.0676 | 0.0688 | 0.1996 |
| | | .1 | .5 | .125 | 0.0768 | 0.076 | 0.2932 | 0.0664 | 0.0616 | 0.3904 |
| | | .2 | 1.0 | .25 | 0.078 | 0.076 | 0.432 | 0.0668 | 0.0624 | 0.6008 |
| C | .2 | .04 | .2 | .05 | 0.0816 | 0.0816 | 0.1804 | 0.0724 | 0.0732 | 0.202 |
| | | .1 | .5 | .125 | 0.0724 | 0.0732 | 0.2928 | 0.0648 | 0.0632 | 0.3908 |
| | | .2 | 1.0 | .25 | 0.0748 | 0.0756 | 0.4324 | 0.0604 | 0.0644 | 0.6028 |
| D | .2 | .04 | .2 | .05 | 0.0976 | 0.096 | 0.2032 | 0.0788 | 0.0796 | 0.228 |
| | | .1 | .5 | .125 | 0.0888 | 0.0864 | 0.3212 | 0.08 | 0.078 | 0.4216 |
| | | .2 | 1.0 | .25 | 0.0892 | 0.0884 | 0.4556 | 0.0732 | 0.0732 | 0.62 |
| E | .2 | .04 | .2 | .05 | 0.1964 | 0.1972 | 0.3256 | 0.1236 | 0.1276 | 0.3092 |
| | | .1 | .5 | .125 | 0.1316 | 0.1308 | 0.3876 | 0.0804 | 0.0772 | 0.436 |
| | | .2 | 1.0 | .25 | 0.122 | 0.1208 | 0.5104 | 0.0788 | 0.0768 | 0.6228 |
| F | .2 | .04 | .2 | .05 | 0.2196 | 0.2204 | 0.3652 | 0.142 | 0.1448 | 0.3468 |
| | | .1 | .5 | .125 | 0.1568 | 0.1564 | 0.4344 | 0.098 | 0.0944 | 0.4756 |
| | | .2 | 1.0 | .25 | 0.1432 | 0.1436 | 0.548 | 0.0884 | 0.0836 | 0.6524 |

Note: The DGP labels are given by A: $(\theta,\rho,\varphi)=(0.5,0.0,0.0)$, B: $(\theta,\rho,\varphi)=(0.5,0.5,0.0)$, C: $(\theta,\rho,\varphi)=(0.5,0.5,0.5)$, D: $(\theta,\rho,\varphi)=(0.8,0.5,0.5)$, E: $(\theta,\rho,\varphi)=(0.8,0.9,0.5)$, and F: $(\theta,\rho,\varphi)=(0.9,0.9,0.9)$.

Table 1.23: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l = 2$, $b_1 = \frac{b}{\lambda}$, $b_2 = \frac{b}{1-\lambda}$, $T = 50$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .5 | .5 | 1.0 | 1.0 | 0.0888 | 0.0856 | 0.552 | 0.0672 | 0.068 | 0.808 |
| | | 1.0 | 2.0 | 2.0 | 0.0888 | 0.0804 | 0.7256 | 0.0644 | 0.0636 | 0.9396 |
| | .4 | .5 | 1.25 | .83 | 0.0828 | 0.0824 | 0.5572 | 0.0724 | 0.0716 | 0.8072 |
| | .2 | .5 | 2.5 | .625 | 0.11 | 0.1116 | 0.6936 | 0.0828 | 0.08 | 0.8828 |
| | | 1.0 | 5.0 | 1.25 | 0.1116 | 0.1032 | 0.8172 | 0.0744 | 0.0628 | 0.9652 |
| B | .5 | .5 | 1.0 | 1.0 | 0.144 | 0.1476 | 0.6348 | 0.0964 | 0.0844 | 0.824 |
| | | 1.0 | 2.0 | 2.0 | 0.144 | 0.1328 | 0.7748 | 0.0832 | 0.062 | 0.9464 |
| | .4 | .5 | 1.25 | .83 | 0.1384 | 0.1316 | 0.636 | 0.0876 | 0.0784 | 0.8236 |
| | .2 | .5 | 2.5 | .625 | 0.15 | 0.1456 | 0.736 | 0.0816 | 0.076 | 0.882 |
| | | 1.0 | 5.0 | 1.25 | 0.1492 | 0.1376 | 0.8504 | 0.0832 | 0.0744 | 0.9648 |
| C | .5 | .5 | 1.0 | 1.0 | 0.1516 | 0.1548 | 0.6508 | 0.09 | 0.0864 | 0.8332 |
| | | 1.0 | 2.0 | 2.0 | 0.1516 | 0.1412 | 0.7912 | 0.0848 | 0.0712 | 0.9528 |
| | .4 | .5 | 1.25 | .83 | 0.1492 | 0.148 | 0.6516 | 0.092 | 0.0844 | 0.8368 |
| | .2 | .5 | 2.5 | .625 | 0.1564 | 0.1596 | 0.7412 | 0.0892 | 0.08 | 0.8752 |
| | | 1.0 | 5.0 | 1.25 | 0.1624 | 0.146 | 0.8592 | 0.086 | 0.082 | 0.9692 |
| D | .5 | .5 | 1.0 | 1.0 | 0.192 | 0.1944 | 0.6788 | 0.1124 | 0.1032 | 0.8592 |
| | | 1.0 | 2.0 | 2.0 | 0.192 | 0.182 | 0.8192 | 0.1012 | 0.0816 | 0.9588 |
| | .4 | .5 | 1.25 | .83 | 0.19 | 0.1852 | 0.694 | 0.1056 | 0.0976 | 0.8504 |
| | .2 | .5 | 2.5 | .625 | 0.1908 | 0.1928 | 0.7792 | 0.102 | 0.0912 | 0.9008 |
| | | 1.0 | 5.0 | 1.25 | 0.1908 | 0.1756 | 0.8756 | 0.0976 | 0.0824 | 0.9788 |
| E | .5 | .5 | 1.0 | 1.0 | 0.3756 | 0.3856 | 0.7872 | 0.2044 | 0.174 | 0.8788 |
| | | 1.0 | 2.0 | 2.0 | 0.3756 | 0.3608 | 0.872 | 0.1724 | 0.132 | 0.964 |
| | .4 | .5 | 1.25 | .83 | 0.354 | 0.354 | 0.7876 | 0.1728 | 0.156 | 0.8896 |
| | .2 | .5 | 2.5 | .625 | 0.25 | 0.2512 | 0.7972 | 0.1108 | 0.0964 | 0.8852 |
| | | 1.0 | 5.0 | 1.25 | 0.2564 | 0.2412 | 0.8752 | 0.1148 | 0.0892 | 0.9692 |
| F | .5 | .5 | 1.0 | 1.0 | 0.4008 | 0.4044 | 0.7988 | 0.2216 | 0.1904 | 0.8956 |
| | | 1.0 | 2.0 | 2.0 | 0.4008 | 0.388 | 0.8816 | 0.1908 | 0.136 | 0.97 |
| | .4 | .5 | 1.25 | .83 | 0.3832 | 0.3796 | 0.8168 | 0.186 | 0.1608 | 0.9032 |
| | .2 | .5 | 2.5 | .625 | 0.2876 | 0.2836 | 0.8148 | 0.1284 | 0.1144 | 0.8984 |
| | | 1.0 | 5.0 | 1.25 | 0.2892 | 0.2756 | 0.8948 | 0.1272 | 0.0972 | 0.974 |

Note: The DGP labels are given by A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$, B: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.0)$, C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$, D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$, E: $(\theta, \rho, \varphi) = (0.8, 0.9, 0.5)$, and F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$.

Table 1.24: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l = 2$, $b_1 = \frac{b}{\lambda}$, $b_2 = \frac{b}{1-\lambda}$, $T = 100$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .5 | .5 | 1.0 | 1.0 | 0.0668 | 0.058 | 0.5344 | 0.0544 | 0.056 | 0.7856 |
| | | 1.0 | 2.0 | 2.0 | 0.0668 | 0.0568 | 0.7052 | 0.052 | 0.054 | 0.9316 |
| | .4 | .5 | 1.25 | .83 | 0.0664 | 0.0664 | 0.5272 | 0.0556 | 0.052 | 0.7896 |
| | .2 | .5 | 2.5 | .625 | 0.0832 | 0.0832 | 0.666 | 0.0688 | 0.0684 | 0.8724 |
| | | 1.0 | 5.0 | 1.25 | 0.0872 | 0.076 | 0.7996 | 0.0676 | 0.0628 | 0.9668 |
| B | .5 | .5 | 1.0 | 1.0 | 0.0856 | 0.088 | 0.5776 | 0.0604 | 0.0636 | 0.804 |
| | | 1.0 | 2.0 | 2.0 | 0.0856 | 0.0784 | 0.7428 | 0.0576 | 0.0508 | 0.9412 |
| | .4 | .5 | 1.25 | .83 | 0.1004 | 0.1008 | 0.5844 | 0.0692 | 0.0668 | 0.8012 |
| | .2 | .5 | 2.5 | .625 | 0.1108 | 0.1092 | 0.6988 | 0.0712 | 0.0716 | 0.8716 |
| | | 1.0 | 5.0 | 1.25 | 0.1144 | 0.1048 | 0.8288 | 0.0724 | 0.0656 | 0.9672 |
| C | .5 | .5 | 1.0 | 1.0 | 0.0988 | 0.0956 | 0.5816 | 0.0648 | 0.0588 | 0.8076 |
| | | 1.0 | 2.0 | 2.0 | 0.0988 | 0.0852 | 0.7492 | 0.0616 | 0.0552 | 0.9428 |
| | .4 | .5 | 1.25 | .83 | 0.1112 | 0.108 | 0.5892 | 0.07 | 0.0696 | 0.8028 |
| | .2 | .5 | 2.5 | .625 | 0.1228 | 0.1228 | 0.7184 | 0.0816 | 0.0772 | 0.872 |
| | | 1.0 | 5.0 | 1.25 | 0.1216 | 0.1076 | 0.8356 | 0.0764 | 0.0664 | 0.97 |
| D | .5 | .5 | 1.0 | 1.0 | 0.1168 | 0.1228 | 0.61 | 0.082 | 0.074 | 0.8096 |
| | | 1.0 | 2.0 | 2.0 | 0.1168 | 0.1104 | 0.7588 | 0.078 | 0.064 | 0.9384 |
| | .4 | .5 | 1.25 | .83 | 0.134 | 0.1312 | 0.6272 | 0.0804 | 0.0804 | 0.8204 |
| | .2 | .5 | 2.5 | .625 | 0.1644 | 0.162 | 0.7584 | 0.0872 | 0.084 | 0.8912 |
| | | 1.0 | 5.0 | 1.25 | 0.166 | 0.152 | 0.856 | 0.0832 | 0.0696 | 0.9696 |
| E | .5 | .5 | 1.0 | 1.0 | 0.2756 | 0.2852 | 0.742 | 0.1412 | 0.1304 | 0.87 |
| | | 1.0 | 2.0 | 2.0 | 0.2756 | 0.2596 | 0.8532 | 0.126 | 0.0932 | 0.9584 |
| | .4 | .5 | 1.25 | .83 | 0.2852 | 0.2792 | 0.7604 | 0.1364 | 0.1276 | 0.8852 |
| | .2 | .5 | 2.5 | .625 | 0.2128 | 0.2112 | 0.7768 | 0.0948 | 0.088 | 0.8852 |
| | | 1.0 | 5.0 | 1.25 | 0.216 | 0.2 | 0.8808 | 0.0992 | 0.0812 | 0.974 |
| F | .5 | .5 | 1.0 | 1.0 | 0.3044 | 0.3028 | 0.7664 | 0.1708 | 0.154 | 0.8764 |
| | | 1.0 | 2.0 | 2.0 | 0.3044 | 0.2872 | 0.8612 | 0.1596 | 0.1192 | 0.96 |
| | .4 | .5 | 1.25 | .83 | 0.2936 | 0.2948 | 0.776 | 0.1528 | 0.136 | 0.884 |
| | .2 | .5 | 2.5 | .625 | 0.2384 | 0.2408 | 0.7892 | 0.1152 | 0.0976 | 0.8864 |
| | | 1.0 | 5.0 | 1.25 | 0.2408 | 0.2252 | 0.8804 | 0.1168 | 0.0872 | 0.9728 |

Note: The DGP labels are given by A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$, B: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.0)$, C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$, D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$, E: $(\theta, \rho, \varphi) = (0.8, 0.9, 0.5)$, and F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$.

Table 1.25: The Finite Sample Size of the Tests Based on $Wald^{(S)}$ and $Wald^{(F)}$,

$l = 2$, $b_1 = \frac{b}{\lambda}$, $b_2 = \frac{b}{1-\lambda}$, $T = 500$, $H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$ at $t = \lambda T$

| | | | | | Bartlett kernel | | | QS kernel | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ | $Wald^{(S)}$ | $Wald^{(F)}$ | $Wald^{(F)}$ |
| DGP | $\lambda$ | $b$ | $b_1$ | $b_2$ | Fixed-$b$ | Fixed-$b$ | chi square | Fixed-$b$ | Fixed-$b$ | chi square |
| A | .5 | .5 | 1.0 | 1.0 | 0.0488 | 0.0532 | 0.4988 | 0.0532 | 0.0496 | 0.7708 |
| | | 1.0 | 2.0 | 2.0 | 0.0488 | 0.0428 | 0.6856 | 0.0528 | 0.0492 | 0.9264 |
| | .4 | .5 | 1.25 | .83 | 0.064 | 0.0612 | 0.5144 | 0.0568 | 0.058 | 0.7848 |
| | .2 | .5 | 2.5 | .625 | 0.0664 | 0.0676 | 0.6488 | 0.056 | 0.058 | 0.8664 |
| | | 1.0 | 5.0 | 1.25 | 0.0636 | 0.056 | 0.7952 | 0.0556 | 0.0568 | 0.958 |
| B | .5 | .5 | 1.0 | 1.0 | 0.06 | 0.0652 | 0.5024 | 0.0548 | 0.0472 | 0.776 |
| | | 1.0 | 2.0 | 2.0 | 0.06 | 0.0524 | 0.6948 | 0.05 | 0.0496 | 0.9312 |
| | .4 | .5 | 1.25 | .83 | 0.0692 | 0.0688 | 0.5148 | 0.0576 | 0.0564 | 0.784 |
| | .2 | .5 | 2.5 | .625 | 0.0752 | 0.0764 | 0.6552 | 0.054 | 0.0592 | 0.8624 |
| | | 1.0 | 5.0 | 1.25 | 0.0724 | 0.0652 | 0.7932 | 0.0564 | 0.0576 | 0.9612 |
| C | .5 | .5 | 1.0 | 1.0 | 0.0596 | 0.0632 | 0.514 | 0.0552 | 0.0488 | 0.7744 |
| | | 1.0 | 2.0 | 2.0 | 0.0596 | 0.0556 | 0.6884 | 0.0492 | 0.0428 | 0.9308 |
| | .4 | .5 | 1.25 | .83 | 0.0668 | 0.0688 | 0.528 | 0.0572 | 0.06 | 0.7868 |
| | .2 | .5 | 2.5 | .625 | 0.076 | 0.076 | 0.656 | 0.0556 | 0.0584 | 0.858 |
| | | 1.0 | 5.0 | 1.25 | 0.0736 | 0.066 | 0.798 | 0.0584 | 0.0516 | 0.9572 |
| D | .5 | .5 | 1.0 | 1.0 | 0.0652 | 0.0644 | 0.5264 | 0.0528 | 0.0528 | 0.7824 |
| | | 1.0 | 2.0 | 2.0 | 0.0652 | 0.0568 | 0.7 | 0.0504 | 0.0464 | 0.9364 |
| | .4 | .5 | 1.25 | .83 | 0.0744 | 0.07 | 0.534 | 0.0604 | 0.0548 | 0.792 |
| | .2 | .5 | 2.5 | .625 | 0.0868 | 0.086 | 0.6704 | 0.066 | 0.0696 | 0.8728 |
| | | 1.0 | 5.0 | 1.25 | 0.0884 | 0.0804 | 0.8148 | 0.0612 | 0.058 | 0.9596 |
| E | .5 | .5 | 1.0 | 1.0 | 0.1032 | 0.1024 | 0.5728 | 0.0724 | 0.0676 | 0.7968 |
| | | 1.0 | 2.0 | 2.0 | 0.1032 | 0.0964 | 0.7288 | 0.0692 | 0.0648 | 0.9304 |
| | .4 | .5 | 1.25 | .83 | 0.11 | 0.1076 | 0.582 | 0.0728 | 0.0652 | 0.794 |
| | .2 | .5 | 2.5 | .625 | 0.1188 | 0.1188 | 0.7032 | 0.0796 | 0.0652 | 0.8704 |
| | | 1.0 | 5.0 | 1.25 | 0.1196 | 0.1056 | 0.8352 | 0.0708 | 0.0624 | 0.9624 |
| F | .5 | .5 | 1.0 | 1.0 | 0.12 | 0.1216 | 0.5904 | 0.0784 | 0.07 | 0.8052 |
| | | 1.0 | 2.0 | 2.0 | 0.12 | 0.1116 | 0.7452 | 0.0688 | 0.06 | 0.932 |
| | .4 | .5 | 1.25 | .83 | 0.1224 | 0.1212 | 0.6112 | 0.0712 | 0.0692 | 0.812 |
| | .2 | .5 | 2.5 | .625 | 0.1416 | 0.1424 | 0.7324 | 0.0868 | 0.0796 | 0.872 |
| | | 1.0 | 5.0 | 1.25 | 0.138 | 0.1288 | 0.8432 | 0.0804 | 0.0732 | 0.9692 |

Note: The DGP labels are given by A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$, B: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.0)$, C: $(\theta, \rho, \varphi) = (0.5, 0.5, 0.5)$, D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$, E: $(\theta, \rho, \varphi) = (0.8, 0.9, 0.5)$, and F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$.

Table 1.26: Fixed-$b$ 95% Critical Values of $Wald^{(F)}$ Unknown Break Date, Bartlett kernel, $l$ =2

| | | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | |
|---|---|---|---|---|---|---|---|---|---|
| $b$ | $SupW$ | $MeanW$ | $ExpW$ | $SupW$ | $MeanW$ | $ExpW$ | $SupW$ | $MeanW$ | $ExpW$ |
| 0.02 | 30.293 | 4.861 | 9.588 | 18.230 | 4.235 | 5.051 | 13.542 | 3.263 | 3.539 |
| 0.04 | 48.447 | 5.9489 | 18.1938 | 26.034 | 4.974 | 8.173 | 16.313 | 3.688 | 4.654 |
| 0.06 | 61.976 | 7.0183 | 24.816 | 33.172 | 5.729 | 11.483 | 19.496 | 4.162 | 5.967 |
| 0.08 | 73.862 | 8.001 | 30.656 | 39.957 | 6.496 | 14.695 | 22.812 | 4.617 | 7.364 |
| 0.1 | 84.848 | 8.973 | 36.109 | 46.263 | 7.278 | 17.653 | 26.323 | 5.146 | 8.998 |
| 0.2 | 138.92 | 14.018 | 63.068 | 76.971 | 11.323 | 32.706 | 46.122 | 8.052 | 18.156 |
| 0.3 | 193.94 | 19.113 | 90.408 | 109.11 | 15.596 | 48.657 | 67.262 | 11.216 | 28.446 |
| 0.4 | 254.14 | 24.443 | 120.71 | 142.31 | 20.009 | 65.120 | 89.241 | 14.464 | 39.161 |
| 0.5 | 313.06 | 29.999 | 149.85 | 176.51 | 24.565 | 82.037 | 111.18 | 17.912 | 49.818 |
| 0.6 | 374.36 | 35.304 | 180.46 | 212.05 | 29.202 | 99.596 | 134.00 | 21.386 | 61.205 |
| 0.7 | 433.71 | 40.902 | 210.22 | 245.66 | 33.625 | 116.32 | 153.93 | 24.666 | 70.991 |
| 0.8 | 491.83 | 46.205 | 239.08 | 279.65 | 38.016 | 133.32 | 173.96 | 27.702 | 81.134 |
| 0.9 | 549.63 | 51.450 | 268.05 | 311.37 | 42.238 | 149.22 | 192.52 | 30.670 | 90.145 |
| 1 | 608.99 | 57.142 | 297.78 | 344.26 | 46.623 | 165.51 | 212.76 | 33.936 | 100.36 |

Table 1.27: Fixed-$b$ 95% Critical Values of $Wald^{(F)}$ Unknown Break Date, QS kernel, $l$ =2

| | | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | |
|---|---|---|---|---|---|---|---|---|---|
| $b$ | $SupW$ | $MeanW$ | $ExpW$ | $SupW$ | $MeanW$ | $ExpW$ | $SupW$ | $MeanW$ | $ExpW$ |
| 0.02 | 64.848 | 5.678 | 26.200 | 24.831 | 4.641 | 7.548 | 15.051 | 3.458 | 4.111 |
| 0.04 | 122.00 | 8.102 | 54.483 | 46.350 | 6.059 | 17.433 | 20.670 | 4.205 | 6.401 |
| 0.06 | 161.74 | 10.617 | 74.329 | 68.158 | 7.630 | 28.148 | 28.305 | 5.060 | 9.666 |
| 0.08 | 207.65 | 13.202 | 97.163 | 91.258 | 9.461 | 39.595 | 38.905 | 6.143 | 14.409 |
| 0.1 | 257.31 | 16.139 | 122.02 | 118.67 | 11.671 | 53.066 | 52.759 | 7.491 | 20.987 |
| 0.2 | 832.93 | 40.501 | 409.56 | 452.33 | 30.155 | 219.29 | 240.65 | 19.924 | 113.55 |
| 0.3 | 3339.8 | 99.975 | 1663.0 | 2055.3 | 77.012 | 1020.8 | 1144.7 | 51.677 | 565.45 |
| 0.4 | 13932 | 239.82 | 6959.4 | 8975.9 | 185.18 | 4481.1 | 4771.4 | 124.22 | 2378.8 |
| 0.5 | 47253 | 537.89 | 23620 | 31752 | 411.53 | 15869 | 16684 | 276.98 | 8334.9 |
| 0.6 | 136211 | 1115.4 | 68099 | 91828 | 850.69 | 45907 | 49492 | 580.43 | 24740 |
| 0.7 | 328737 | 2170.5 | 164361 | 224463 | 1674.7 | 112225 | 128234 | 1140.0 | 64110 |
| 0.8 | 719812 | 3982.4 | 359899 | 488008 | 3100.4 | 243997 | 283267 | 2099.3 | 141627 |
| 0.9 | 1444833 | 7015.5 | 722409 | 970172 | 5395.5 | 485079 | 565285 | 3626.6 | 282635 |
| 1 | 2647520 | 11566 | 1323754 | 1829406 | 9072.3 | 914696 | 1062685 | 5951.4 | 531336 |

Table 1.28: The Finite Sample Size of the $SupW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.331 | 0.184 | 0.721 | 0.160 | 0.111 | 0.347 | 0.104 | 0.099 | 0.163 |
| 0.04 | 0.290 | 0.191 | 0.847 | 0.136 | 0.096 | 0.494 | 0.108 | 0.098 | 0.241 |
| 0.06 | 0.284 | 0.212 | 0.898 | 0.131 | 0.104 | 0.601 | 0.103 | 0.082 | 0.308 |
| 0.08 | 0.284 | 0.211 | 0.936 | 0.127 | 0.108 | 0.696 | 0.098 | 0.078 | 0.384 |
| 0.1 | 0.277 | 0.212 | 0.954 | 0.124 | 0.106 | 0.756 | 0.094 | 0.071 | 0.447 |
| 0.2 | 0.262 | 0.153 | 0.992 | 0.125 | 0.081 | 0.908 | 0.084 | 0.056 | 0.678 |
| 0.3 | 0.264 | 0.084 | 0.998 | 0.124 | 0.050 | 0.962 | 0.081 | 0.043 | 0.812 |
| 0.4 | 0.259 | 0.048 | 1.000 | 0.123 | 0.031 | 0.983 | 0.081 | 0.028 | 0.885 |
| 0.5 | 0.250 | 0.036 | 1.000 | 0.120 | 0.023 | 0.994 | 0.081 | 0.019 | 0.923 |
| 0.6 | 0.250 | 0.026 | 1.000 | 0.118 | 0.017 | 0.997 | 0.080 | 0.017 | 0.950 |
| 0.7 | 0.250 | 0.024 | 1.000 | 0.120 | 0.016 | 0.999 | 0.080 | 0.015 | 0.968 |
| 0.8 | 0.243 | 0.020 | 1.000 | 0.119 | 0.014 | 1.000 | 0.079 | 0.012 | 0.984 |
| 0.9 | 0.248 | 0.015 | 1.000 | 0.123 | 0.010 | 1.000 | 0.079 | 0.010 | 0.989 |
| 1 | 0.253 | 0.012 | 1.000 | 0.121 | 0.007 | 1.000 | 0.080 | 0.009 | 0.994 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.093 | 0.084 | 0.472 | 0.069 | 0.070 | 0.217 | 0.062 | 0.062 | 0.111 |
| 0.04 | 0.086 | 0.080 | 0.704 | 0.069 | 0.063 | 0.376 | 0.060 | 0.060 | 0.179 |
| 0.06 | 0.086 | 0.082 | 0.810 | 0.069 | 0.063 | 0.507 | 0.057 | 0.058 | 0.247 |
| 0.08 | 0.087 | 0.079 | 0.865 | 0.065 | 0.058 | 0.607 | 0.060 | 0.056 | 0.315 |
| 0.1 | 0.084 | 0.077 | 0.904 | 0.064 | 0.057 | 0.679 | 0.062 | 0.056 | 0.381 |
| 0.2 | 0.078 | 0.063 | 0.983 | 0.058 | 0.052 | 0.878 | 0.054 | 0.044 | 0.641 |
| 0.3 | 0.081 | 0.055 | 0.995 | 0.057 | 0.052 | 0.946 | 0.061 | 0.052 | 0.786 |
| 0.4 | 0.078 | 0.047 | 0.999 | 0.056 | 0.041 | 0.974 | 0.054 | 0.050 | 0.865 |
| 0.5 | 0.083 | 0.046 | 1.000 | 0.057 | 0.038 | 0.986 | 0.052 | 0.042 | 0.908 |
| 0.6 | 0.081 | 0.034 | 1.000 | 0.059 | 0.034 | 0.994 | 0.050 | 0.041 | 0.935 |
| 0.7 | 0.080 | 0.026 | 1.000 | 0.062 | 0.028 | 0.998 | 0.055 | 0.035 | 0.959 |
| 0.8 | 0.078 | 0.025 | 1.000 | 0.058 | 0.027 | 1.000 | 0.054 | 0.031 | 0.977 |
| 0.9 | 0.079 | 0.026 | 1.000 | 0.058 | 0.025 | 1.000 | 0.050 | 0.030 | 0.987 |
| 1 | 0.081 | 0.026 | 1.000 | 0.061 | 0.024 | 1.000 | 0.055 | 0.028 | 0.992 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.0772 | 0.0784 | 0.4132 | 0.056 | 0.0512 | 0.1932 | 0.0592 | 0.0572 | 0.116 |
| 0.04 | 0.078 | 0.072 | 0.6584 | 0.0544 | 0.0536 | 0.3664 | 0.0592 | 0.0592 | 0.1808 |
| 0.06 | 0.0784 | 0.0712 | 0.7832 | 0.05 | 0.052 | 0.4984 | 0.0572 | 0.0584 | 0.2452 |
| 0.08 | 0.0756 | 0.0592 | 0.8528 | 0.0464 | 0.0484 | 0.5956 | 0.056 | 0.054 | 0.3048 |
| 0.1 | 0.0696 | 0.0664 | 0.8956 | 0.0496 | 0.0532 | 0.6788 | 0.0564 | 0.0516 | 0.3712 |
| 0.2 | 0.0704 | 0.0584 | 0.9804 | 0.05 | 0.0556 | 0.866 | 0.0492 | 0.0472 | 0.6172 |
| 0.3 | 0.07 | 0.0472 | 0.994 | 0.0488 | 0.0476 | 0.934 | 0.0488 | 0.0436 | 0.7744 |
| 0.4 | 0.0696 | 0.0448 | 0.9984 | 0.0432 | 0.0464 | 0.9688 | 0.05 | 0.05 | 0.8536 |
| 0.5 | 0.0716 | 0.0468 | 1 | 0.052 | 0.0452 | 0.9852 | 0.054 | 0.0548 | 0.8964 |
| 0.6 | 0.072 | 0.0456 | 1 | 0.0472 | 0.0484 | 0.9912 | 0.0464 | 0.05 | 0.928 |
| 0.7 | 0.0724 | 0.0472 | 1 | 0.046 | 0.0492 | 0.996 | 0.0496 | 0.0468 | 0.956 |
| 0.8 | 0.0664 | 0.0436 | 1 | 0.054 | 0.0484 | 0.9992 | 0.0472 | 0.0468 | 0.974 |
| 0.9 | 0.0696 | 0.04 | 1 | 0.0508 | 0.0456 | 0.9996 | 0.0484 | 0.0456 | 0.9852 |
| 1 | 0.0708 | 0.04 | 1 | 0.0524 | 0.044 | 1 | 0.0484 | 0.0424 | 0.9928 |

Table 1.29: The Finite Sample Size of the $SupW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.665 | 0.308 | 0.947 | 0.598 | 0.395 | 0.801 | 0.493 | 0.376 | 0.604 |
| 0.04 | 0.466 | 0.173 | 0.955 | 0.417 | 0.200 | 0.800 | 0.341 | 0.227 | 0.542 |
| 0.06 | 0.398 | 0.166 | 0.966 | 0.355 | 0.166 | 0.842 | 0.292 | 0.172 | 0.587 |
| 0.08 | 0.365 | 0.159 | 0.978 | 0.319 | 0.156 | 0.881 | 0.268 | 0.145 | 0.638 |
| 0.1 | 0.351 | 0.161 | 0.985 | 0.308 | 0.153 | 0.906 | 0.255 | 0.128 | 0.674 |
| 0.2 | 0.316 | 0.114 | 0.996 | 0.283 | 0.101 | 0.968 | 0.224 | 0.077 | 0.832 |
| 0.3 | 0.308 | 0.071 | 1.000 | 0.271 | 0.061 | 0.990 | 0.216 | 0.054 | 0.907 |
| 0.4 | 0.304 | 0.044 | 1.000 | 0.266 | 0.036 | 0.997 | 0.213 | 0.037 | 0.948 |
| 0.5 | 0.302 | 0.034 | 1.000 | 0.272 | 0.026 | 0.999 | 0.213 | 0.034 | 0.966 |
| 0.6 | 0.310 | 0.024 | 1.000 | 0.276 | 0.023 | 1.000 | 0.208 | 0.028 | 0.980 |
| 0.7 | 0.310 | 0.020 | 1.000 | 0.266 | 0.020 | 1.000 | 0.210 | 0.024 | 0.989 |
| 0.8 | 0.307 | 0.021 | 1.000 | 0.265 | 0.021 | 1.000 | 0.211 | 0.020 | 0.993 |
| 0.9 | 0.304 | 0.021 | 1.000 | 0.265 | 0.020 | 1.000 | 0.209 | 0.020 | 0.995 |
| 1 | 0.309 | 0.020 | 1.000 | 0.273 | 0.021 | 1.000 | 0.209 | 0.018 | 0.999 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.247 | 0.109 | 0.697 | 0.193 | 0.112 | 0.391 | 0.131 | 0.090 | 0.216 |
| 0.04 | 0.207 | 0.097 | 0.832 | 0.149 | 0.085 | 0.514 | 0.108 | 0.084 | 0.251 |
| 0.06 | 0.186 | 0.096 | 0.888 | 0.139 | 0.083 | 0.616 | 0.102 | 0.077 | 0.318 |
| 0.08 | 0.181 | 0.086 | 0.930 | 0.126 | 0.078 | 0.685 | 0.098 | 0.070 | 0.382 |
| 0.1 | 0.170 | 0.089 | 0.956 | 0.129 | 0.077 | 0.756 | 0.096 | 0.070 | 0.447 |
| 0.2 | 0.144 | 0.074 | 0.991 | 0.114 | 0.058 | 0.908 | 0.088 | 0.051 | 0.684 |
| 0.3 | 0.146 | 0.054 | 0.998 | 0.110 | 0.050 | 0.964 | 0.089 | 0.044 | 0.824 |
| 0.4 | 0.140 | 0.036 | 0.999 | 0.106 | 0.036 | 0.985 | 0.088 | 0.041 | 0.888 |
| 0.5 | 0.144 | 0.037 | 1.000 | 0.107 | 0.034 | 0.993 | 0.083 | 0.041 | 0.926 |
| 0.6 | 0.145 | 0.032 | 1.000 | 0.105 | 0.038 | 0.998 | 0.083 | 0.038 | 0.949 |
| 0.7 | 0.146 | 0.031 | 1.000 | 0.103 | 0.033 | 0.999 | 0.087 | 0.036 | 0.970 |
| 0.8 | 0.144 | 0.032 | 1.000 | 0.106 | 0.034 | 1.000 | 0.090 | 0.033 | 0.984 |
| 0.9 | 0.146 | 0.031 | 1.000 | 0.106 | 0.030 | 1.000 | 0.087 | 0.030 | 0.990 |
| 1 | 0.149 | 0.032 | 1.000 | 0.108 | 0.031 | 1.000 | 0.087 | 0.030 | 0.994 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews | Fixed-*b* | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.1628 | 0.0928 | 0.5716 | 0.1176 | 0.0836 | 0.2892 | 0.0952 | 0.0744 | 0.1596 |
| 0.04 | 0.144 | 0.0888 | 0.7504 | 0.0996 | 0.0636 | 0.4304 | 0.0884 | 0.0732 | 0.2132 |
| 0.06 | 0.1344 | 0.0844 | 0.844 | 0.0924 | 0.0576 | 0.5604 | 0.0836 | 0.0732 | 0.2776 |
| 0.08 | 0.1308 | 0.0796 | 0.8956 | 0.0856 | 0.058 | 0.654 | 0.0828 | 0.066 | 0.3492 |
| 0.1 | 0.122 | 0.0672 | 0.9248 | 0.0864 | 0.056 | 0.724 | 0.0816 | 0.0592 | 0.4104 |
| 0.2 | 0.1196 | 0.0648 | 0.9828 | 0.0828 | 0.052 | 0.8888 | 0.0624 | 0.0504 | 0.6548 |
| 0.3 | 0.1044 | 0.0512 | 0.9964 | 0.0752 | 0.058 | 0.9516 | 0.0648 | 0.0544 | 0.7856 |
| 0.4 | 0.1072 | 0.0436 | 0.9996 | 0.0768 | 0.046 | 0.9748 | 0.0664 | 0.0516 | 0.862 |
| 0.5 | 0.1056 | 0.0424 | 0.9996 | 0.076 | 0.04 | 0.9884 | 0.0648 | 0.0512 | 0.9076 |
| 0.6 | 0.1056 | 0.042 | 1 | 0.0772 | 0.0392 | 0.9952 | 0.0692 | 0.0436 | 0.936 |
| 0.7 | 0.1084 | 0.0412 | 1 | 0.0804 | 0.0392 | 0.9992 | 0.0688 | 0.0412 | 0.9568 |
| 0.8 | 0.1048 | 0.0416 | 1 | 0.0804 | 0.0404 | 0.9992 | 0.066 | 0.0408 | 0.9788 |
| 0.9 | 0.1068 | 0.0404 | 1 | 0.082 | 0.0404 | 1 | 0.0672 | 0.0364 | 0.9884 |
| 1 | 0.1088 | 0.0404 | 1 | 0.0788 | 0.0404 | 1 | 0.0668 | 0.0396 | 0.9912 |

Table 1.30: The Finite Sample Size of the $SupW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.934 | 0.586 | 0.998 | 0.963 | 0.879 | 0.992 | 0.942 | 0.886 | 0.967 |
| 0.04 | 0.672 | 0.198 | 0.994 | 0.806 | 0.472 | 0.976 | 0.810 | 0.652 | 0.914 |
| 0.06 | 0.508 | 0.136 | 0.991 | 0.680 | 0.293 | 0.971 | 0.700 | 0.487 | 0.895 |
| 0.08 | 0.405 | 0.100 | 0.991 | 0.578 | 0.227 | 0.973 | 0.632 | 0.369 | 0.895 |
| 0.1 | 0.346 | 0.092 | 0.990 | 0.528 | 0.184 | 0.974 | 0.596 | 0.304 | 0.900 |
| 0.2 | 0.272 | 0.073 | 0.994 | 0.445 | 0.122 | 0.986 | 0.507 | 0.160 | 0.948 |
| 0.3 | 0.286 | 0.060 | 0.998 | 0.440 | 0.081 | 0.995 | 0.503 | 0.104 | 0.972 |
| 0.4 | 0.299 | 0.045 | 0.999 | 0.444 | 0.057 | 0.998 | 0.492 | 0.082 | 0.986 |
| 0.5 | 0.307 | 0.036 | 1.000 | 0.447 | 0.046 | 1.000 | 0.488 | 0.064 | 0.994 |
| 0.6 | 0.300 | 0.034 | 1.000 | 0.448 | 0.038 | 1.000 | 0.498 | 0.052 | 0.997 |
| 0.7 | 0.296 | 0.033 | 1.000 | 0.442 | 0.036 | 1.000 | 0.499 | 0.042 | 0.999 |
| 0.8 | 0.291 | 0.030 | 1.000 | 0.442 | 0.034 | 1.000 | 0.492 | 0.038 | 0.999 |
| 0.9 | 0.290 | 0.026 | 1.000 | 0.443 | 0.033 | 1.000 | 0.492 | 0.035 | 1.000 |
| 1 | 0.296 | 0.026 | 1.000 | 0.446 | 0.030 | 1.000 | 0.494 | 0.032 | 1.000 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.591 | 0.195 | 0.945 | 0.608 | 0.351 | 0.812 | 0.512 | 0.354 | 0.632 |
| 0.04 | 0.347 | 0.093 | 0.941 | 0.385 | 0.142 | 0.779 | 0.313 | 0.190 | 0.527 |
| 0.06 | 0.276 | 0.075 | 0.950 | 0.306 | 0.107 | 0.812 | 0.252 | 0.139 | 0.545 |
| 0.08 | 0.239 | 0.070 | 0.962 | 0.260 | 0.090 | 0.846 | 0.227 | 0.114 | 0.587 |
| 0.1 | 0.212 | 0.059 | 0.971 | 0.239 | 0.083 | 0.874 | 0.204 | 0.098 | 0.630 |
| 0.2 | 0.184 | 0.054 | 0.990 | 0.208 | 0.062 | 0.952 | 0.177 | 0.066 | 0.802 |
| 0.3 | 0.178 | 0.050 | 0.998 | 0.202 | 0.056 | 0.983 | 0.183 | 0.062 | 0.878 |
| 0.4 | 0.180 | 0.040 | 1.000 | 0.200 | 0.050 | 0.994 | 0.184 | 0.059 | 0.930 |
| 0.5 | 0.176 | 0.041 | 1.000 | 0.201 | 0.043 | 0.998 | 0.180 | 0.050 | 0.956 |
| 0.6 | 0.180 | 0.039 | 1.000 | 0.192 | 0.042 | 1.000 | 0.178 | 0.048 | 0.973 |
| 0.7 | 0.176 | 0.042 | 1.000 | 0.198 | 0.044 | 1.000 | 0.176 | 0.044 | 0.987 |
| 0.8 | 0.170 | 0.038 | 1.000 | 0.197 | 0.042 | 1.000 | 0.179 | 0.043 | 0.993 |
| 0.9 | 0.173 | 0.038 | 1.000 | 0.198 | 0.040 | 1.000 | 0.174 | 0.043 | 0.995 |
| 1 | 0.174 | 0.038 | 1.000 | 0.200 | 0.039 | 1.000 | 0.174 | 0.044 | 0.997 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews | Fixed-$b$ | | Andrews |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.4004 | 0.1216 | 0.8364 | 0.3504 | 0.1736 | 0.5964 | 0.2788 | 0.1832 | 0.3852 |
| 0.04 | 0.2688 | 0.0724 | 0.888 | 0.2204 | 0.0892 | 0.6408 | 0.192 | 0.1188 | 0.3712 |
| 0.06 | 0.2176 | 0.0688 | 0.9216 | 0.1764 | 0.068 | 0.7052 | 0.156 | 0.0928 | 0.4092 |
| 0.08 | 0.1956 | 0.0564 | 0.9436 | 0.1628 | 0.0576 | 0.772 | 0.1492 | 0.0812 | 0.478 |
| 0.1 | 0.1728 | 0.0516 | 0.9568 | 0.1504 | 0.054 | 0.8152 | 0.1392 | 0.0696 | 0.5344 |
| 0.2 | 0.1504 | 0.046 | 0.9908 | 0.132 | 0.0528 | 0.9236 | 0.1232 | 0.0476 | 0.7312 |
| 0.3 | 0.1464 | 0.0444 | 0.9964 | 0.1332 | 0.0448 | 0.9712 | 0.118 | 0.0452 | 0.8332 |
| 0.4 | 0.1464 | 0.0424 | 0.9996 | 0.132 | 0.0352 | 0.9864 | 0.122 | 0.036 | 0.8908 |
| 0.5 | 0.1532 | 0.04 | 0.9996 | 0.1288 | 0.0352 | 0.9924 | 0.1232 | 0.036 | 0.924 |
| 0.6 | 0.1556 | 0.0396 | 0.9996 | 0.1364 | 0.0388 | 0.9968 | 0.126 | 0.038 | 0.9536 |
| 0.7 | 0.1532 | 0.0408 | 1 | 0.1332 | 0.0392 | 0.9992 | 0.1164 | 0.0376 | 0.9736 |
| 0.8 | 0.1416 | 0.042 | 1 | 0.1372 | 0.0412 | 0.9996 | 0.1192 | 0.0388 | 0.9856 |
| 0.9 | 0.1452 | 0.042 | 1 | 0.1292 | 0.0424 | 0.9996 | 0.1204 | 0.0388 | 0.9936 |
| 1 | 0.1448 | 0.0444 | 1 | 0.1296 | 0.0408 | 1 | 0.1244 | 0.04 | 0.9968 |

Table 1.31: The Finite Sample Size of the $MeanW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.162 | 0.148 | 0.290 | 0.101 | 0.094 | 0.157 | 0.084 | 0.082 | 0.120 |
| 0.04 | 0.175 | 0.190 | 0.446 | 0.105 | 0.105 | 0.243 | 0.089 | 0.090 | 0.156 |
| 0.06 | 0.182 | 0.202 | 0.586 | 0.108 | 0.114 | 0.328 | 0.090 | 0.087 | 0.201 |
| 0.08 | 0.192 | 0.215 | 0.686 | 0.107 | 0.114 | 0.406 | 0.090 | 0.088 | 0.243 |
| 0.1 | 0.190 | 0.226 | 0.759 | 0.103 | 0.116 | 0.482 | 0.086 | 0.081 | 0.291 |
| 0.2 | 0.204 | 0.194 | 0.941 | 0.109 | 0.109 | 0.750 | 0.086 | 0.079 | 0.518 |
| 0.3 | 0.209 | 0.165 | 0.984 | 0.114 | 0.096 | 0.895 | 0.081 | 0.068 | 0.674 |
| 0.4 | 0.208 | 0.132 | 0.996 | 0.117 | 0.087 | 0.950 | 0.086 | 0.069 | 0.771 |
| 0.5 | 0.216 | 0.120 | 0.999 | 0.119 | 0.080 | 0.977 | 0.089 | 0.065 | 0.850 |
| 0.6 | 0.212 | 0.112 | 1.000 | 0.121 | 0.080 | 0.991 | 0.087 | 0.058 | 0.902 |
| 0.7 | 0.216 | 0.110 | 1.000 | 0.120 | 0.075 | 0.996 | 0.084 | 0.059 | 0.931 |
| 0.8 | 0.216 | 0.102 | 1.000 | 0.123 | 0.071 | 1.000 | 0.086 | 0.058 | 0.955 |
| 0.9 | 0.218 | 0.097 | 1.000 | 0.121 | 0.068 | 1.000 | 0.088 | 0.056 | 0.971 |
| 1 | 0.217 | 0.097 | 1.000 | 0.122 | 0.066 | 1.000 | 0.087 | 0.055 | 0.982 |
| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.064 | 0.066 | 0.134 | 0.063 | 0.060 | 0.103 | 0.056 | 0.060 | 0.084 |
| 0.04 | 0.063 | 0.066 | 0.248 | 0.058 | 0.060 | 0.163 | 0.054 | 0.056 | 0.113 |
| 0.06 | 0.060 | 0.069 | 0.363 | 0.058 | 0.064 | 0.235 | 0.056 | 0.058 | 0.151 |
| 0.08 | 0.063 | 0.068 | 0.471 | 0.059 | 0.061 | 0.302 | 0.059 | 0.056 | 0.189 |
| 0.1 | 0.064 | 0.070 | 0.570 | 0.059 | 0.054 | 0.372 | 0.061 | 0.055 | 0.231 |
| 0.2 | 0.069 | 0.066 | 0.853 | 0.056 | 0.054 | 0.666 | 0.056 | 0.056 | 0.465 |
| 0.3 | 0.071 | 0.065 | 0.956 | 0.056 | 0.057 | 0.827 | 0.055 | 0.060 | 0.628 |
| 0.4 | 0.068 | 0.057 | 0.986 | 0.055 | 0.053 | 0.911 | 0.062 | 0.055 | 0.741 |
| 0.5 | 0.065 | 0.057 | 0.995 | 0.054 | 0.052 | 0.954 | 0.060 | 0.051 | 0.818 |
| 0.6 | 0.065 | 0.056 | 1.000 | 0.059 | 0.055 | 0.975 | 0.054 | 0.052 | 0.869 |
| 0.7 | 0.068 | 0.058 | 1.000 | 0.057 | 0.052 | 0.990 | 0.057 | 0.051 | 0.912 |
| 0.8 | 0.069 | 0.058 | 1.000 | 0.058 | 0.048 | 0.996 | 0.056 | 0.050 | 0.940 |
| 0.9 | 0.068 | 0.055 | 1.000 | 0.058 | 0.047 | 0.998 | 0.056 | 0.049 | 0.961 |
| 1 | 0.067 | 0.055 | 1.000 | 0.054 | 0.050 | 1.000 | 0.060 | 0.050 | 0.971 |
| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.0576 | 0.0576 | 0.1296 | 0.0508 | 0.0524 | 0.0988 | 0.0468 | 0.0476 | 0.0812 |
| 0.04 | 0.0604 | 0.0616 | 0.224 | 0.0524 | 0.0532 | 0.1516 | 0.0476 | 0.05 | 0.112 |
| 0.06 | 0.0596 | 0.058 | 0.3244 | 0.0516 | 0.0484 | 0.2188 | 0.0496 | 0.0484 | 0.1508 |
| 0.08 | 0.0588 | 0.0584 | 0.4352 | 0.0504 | 0.05 | 0.2896 | 0.0504 | 0.05 | 0.1904 |
| 0.1 | 0.0576 | 0.06 | 0.5264 | 0.0504 | 0.0484 | 0.352 | 0.0512 | 0.0464 | 0.2332 |
| 0.2 | 0.0556 | 0.0588 | 0.8344 | 0.052 | 0.0488 | 0.6536 | 0.048 | 0.0428 | 0.4424 |
| 0.3 | 0.0556 | 0.0512 | 0.942 | 0.0492 | 0.046 | 0.8236 | 0.0508 | 0.0492 | 0.6132 |
| 0.4 | 0.0532 | 0.0536 | 0.974 | 0.0472 | 0.0472 | 0.9012 | 0.048 | 0.0496 | 0.736 |
| 0.5 | 0.0592 | 0.056 | 0.9912 | 0.0544 | 0.0508 | 0.9424 | 0.0504 | 0.052 | 0.8072 |
| 0.6 | 0.0584 | 0.0516 | 0.9976 | 0.052 | 0.0508 | 0.9696 | 0.0476 | 0.0452 | 0.8616 |
| 0.7 | 0.056 | 0.0492 | 0.9988 | 0.0512 | 0.0504 | 0.9868 | 0.0488 | 0.0432 | 0.9028 |
| 0.8 | 0.058 | 0.0516 | 1 | 0.0496 | 0.048 | 0.9956 | 0.0472 | 0.0476 | 0.9328 |
| 0.9 | 0.0556 | 0.0516 | 1 | 0.0516 | 0.05 | 0.9984 | 0.0492 | 0.0472 | 0.958 |
| 1 | 0.056 | 0.0528 | 1 | 0.0512 | 0.0468 | 1 | 0.0504 | 0.0476 | 0.9712 |

Table 1.32: The Finite Sample Size of the $MeanW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.664 | 0.530 | 0.806 | 0.527 | 0.420 | 0.632 | 0.420 | 0.324 | 0.488 |
| 0.04 | 0.523 | 0.365 | 0.806 | 0.386 | 0.273 | 0.593 | 0.283 | 0.198 | 0.414 |
| 0.06 | 0.476 | 0.312 | 0.852 | 0.343 | 0.245 | 0.646 | 0.235 | 0.167 | 0.431 |
| 0.08 | 0.456 | 0.301 | 0.893 | 0.328 | 0.231 | 0.700 | 0.229 | 0.158 | 0.475 |
| 0.1 | 0.445 | 0.291 | 0.926 | 0.317 | 0.219 | 0.752 | 0.216 | 0.154 | 0.521 |
| 0.2 | 0.432 | 0.234 | 0.985 | 0.310 | 0.176 | 0.914 | 0.208 | 0.136 | 0.712 |
| 0.3 | 0.440 | 0.201 | 0.996 | 0.325 | 0.156 | 0.973 | 0.214 | 0.121 | 0.822 |
| 0.4 | 0.448 | 0.178 | 1.000 | 0.333 | 0.138 | 0.990 | 0.223 | 0.118 | 0.897 |
| 0.5 | 0.462 | 0.155 | 1.000 | 0.347 | 0.127 | 0.997 | 0.224 | 0.114 | 0.938 |
| 0.6 | 0.457 | 0.146 | 1.000 | 0.342 | 0.130 | 0.998 | 0.224 | 0.106 | 0.958 |
| 0.7 | 0.460 | 0.143 | 1.000 | 0.339 | 0.122 | 1.000 | 0.222 | 0.099 | 0.977 |
| 0.8 | 0.460 | 0.134 | 1.000 | 0.342 | 0.115 | 1.000 | 0.226 | 0.094 | 0.984 |
| 0.9 | 0.456 | 0.129 | 1.000 | 0.339 | 0.111 | 1.000 | 0.226 | 0.090 | 0.992 |
| 1 | 0.455 | 0.121 | 1.000 | 0.341 | 0.105 | 1.000 | 0.229 | 0.089 | 0.996 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.160 | 0.110 | 0.291 | 0.127 | 0.091 | 0.201 | 0.109 | 0.085 | 0.152 |
| 0.04 | 0.142 | 0.103 | 0.412 | 0.110 | 0.084 | 0.255 | 0.089 | 0.070 | 0.172 |
| 0.06 | 0.133 | 0.106 | 0.519 | 0.103 | 0.087 | 0.330 | 0.086 | 0.070 | 0.214 |
| 0.08 | 0.134 | 0.102 | 0.614 | 0.101 | 0.088 | 0.396 | 0.083 | 0.068 | 0.254 |
| 0.1 | 0.128 | 0.103 | 0.694 | 0.099 | 0.084 | 0.468 | 0.081 | 0.070 | 0.295 |
| 0.2 | 0.136 | 0.086 | 0.915 | 0.103 | 0.074 | 0.737 | 0.080 | 0.061 | 0.502 |
| 0.3 | 0.136 | 0.088 | 0.976 | 0.105 | 0.077 | 0.879 | 0.082 | 0.063 | 0.675 |
| 0.4 | 0.139 | 0.075 | 0.994 | 0.101 | 0.071 | 0.941 | 0.084 | 0.061 | 0.776 |
| 0.5 | 0.148 | 0.078 | 0.999 | 0.104 | 0.072 | 0.974 | 0.082 | 0.065 | 0.848 |
| 0.6 | 0.147 | 0.078 | 1.000 | 0.106 | 0.072 | 0.989 | 0.083 | 0.062 | 0.900 |
| 0.7 | 0.144 | 0.077 | 1.000 | 0.105 | 0.070 | 0.996 | 0.085 | 0.062 | 0.929 |
| 0.8 | 0.144 | 0.075 | 1.000 | 0.108 | 0.068 | 0.999 | 0.084 | 0.061 | 0.955 |
| 0.9 | 0.144 | 0.074 | 1.000 | 0.104 | 0.070 | 1.000 | 0.084 | 0.062 | 0.970 |
| 1 | 0.141 | 0.073 | 1.000 | 0.105 | 0.067 | 1.000 | 0.086 | 0.065 | 0.982 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.1044 | 0.082 | 0.202 | 0.086 | 0.0672 | 0.1416 | 0.0732 | 0.0612 | 0.1096 |
| 0.04 | 0.0932 | 0.0784 | 0.3076 | 0.0768 | 0.0632 | 0.1972 | 0.0696 | 0.0628 | 0.1332 |
| 0.06 | 0.0952 | 0.0732 | 0.434 | 0.0768 | 0.0624 | 0.2692 | 0.066 | 0.0552 | 0.1768 |
| 0.08 | 0.0928 | 0.0732 | 0.5352 | 0.0716 | 0.0588 | 0.3376 | 0.0628 | 0.0544 | 0.2144 |
| 0.1 | 0.0912 | 0.0696 | 0.6228 | 0.0676 | 0.0624 | 0.4068 | 0.0608 | 0.0528 | 0.2552 |
| 0.2 | 0.094 | 0.068 | 0.8792 | 0.0716 | 0.0636 | 0.6992 | 0.0588 | 0.0536 | 0.464 |
| 0.3 | 0.0952 | 0.0688 | 0.9568 | 0.072 | 0.0624 | 0.8524 | 0.0608 | 0.062 | 0.6356 |
| 0.4 | 0.0956 | 0.0648 | 0.9844 | 0.0744 | 0.0624 | 0.9208 | 0.0664 | 0.0596 | 0.7512 |
| 0.5 | 0.0948 | 0.0616 | 0.996 | 0.0736 | 0.054 | 0.9552 | 0.0612 | 0.0616 | 0.822 |
| 0.6 | 0.0984 | 0.058 | 0.9984 | 0.0736 | 0.0532 | 0.9772 | 0.0624 | 0.056 | 0.8664 |
| 0.7 | 0.0968 | 0.0608 | 1 | 0.072 | 0.0536 | 0.9912 | 0.0604 | 0.0552 | 0.9092 |
| 0.8 | 0.0984 | 0.0604 | 1 | 0.0728 | 0.0532 | 0.9972 | 0.0592 | 0.054 | 0.938 |
| 0.9 | 0.0952 | 0.0616 | 1 | 0.072 | 0.054 | 0.9992 | 0.0616 | 0.0548 | 0.9604 |
| 1 | 0.096 | 0.062 | 1 | 0.0716 | 0.0512 | 1 | 0.0628 | 0.0568 | 0.974 |

Table 1.33: The Finite Sample Size of the $MeanW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.985 | 0.948 | 0.996 | 0.964 | 0.924 | 0.982 | 0.916 | 0.854 | 0.943 |
| 0.04 | 0.910 | 0.758 | 0.985 | 0.859 | 0.735 | 0.945 | 0.772 | 0.658 | 0.858 |
| 0.06 | 0.825 | 0.591 | 0.979 | 0.783 | 0.610 | 0.932 | 0.680 | 0.538 | 0.824 |
| 0.08 | 0.771 | 0.478 | 0.978 | 0.724 | 0.513 | 0.937 | 0.627 | 0.462 | 0.822 |
| 0.1 | 0.723 | 0.407 | 0.979 | 0.687 | 0.448 | 0.942 | 0.586 | 0.409 | 0.838 |
| 0.2 | 0.650 | 0.269 | 0.992 | 0.628 | 0.302 | 0.976 | 0.533 | 0.302 | 0.912 |
| 0.3 | 0.646 | 0.243 | 0.997 | 0.619 | 0.260 | 0.991 | 0.543 | 0.252 | 0.955 |
| 0.4 | 0.655 | 0.218 | 0.999 | 0.634 | 0.233 | 0.997 | 0.545 | 0.235 | 0.975 |
| 0.5 | 0.668 | 0.207 | 1.000 | 0.641 | 0.228 | 0.998 | 0.552 | 0.223 | 0.988 |
| 0.6 | 0.670 | 0.201 | 1.000 | 0.642 | 0.215 | 1.000 | 0.546 | 0.207 | 0.996 |
| 0.7 | 0.659 | 0.194 | 1.000 | 0.638 | 0.204 | 1.000 | 0.537 | 0.194 | 0.998 |
| 0.8 | 0.659 | 0.184 | 1.000 | 0.631 | 0.195 | 1.000 | 0.546 | 0.180 | 0.998 |
| 0.9 | 0.662 | 0.182 | 1.000 | 0.640 | 0.189 | 1.000 | 0.546 | 0.175 | 0.998 |
| 1 | 0.664 | 0.174 | 1.000 | 0.641 | 0.181 | 1.000 | 0.547 | 0.171 | 0.999 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.653 | 0.451 | 0.788 | 0.532 | 0.390 | 0.645 | 0.429 | 0.318 | 0.502 |
| 0.04 | 0.458 | 0.241 | 0.761 | 0.360 | 0.225 | 0.568 | 0.284 | 0.193 | 0.404 |
| 0.06 | 0.382 | 0.181 | 0.794 | 0.302 | 0.179 | 0.595 | 0.235 | 0.157 | 0.408 |
| 0.08 | 0.342 | 0.160 | 0.833 | 0.275 | 0.156 | 0.642 | 0.217 | 0.137 | 0.432 |
| 0.1 | 0.324 | 0.145 | 0.868 | 0.261 | 0.140 | 0.691 | 0.206 | 0.124 | 0.469 |
| 0.2 | 0.296 | 0.118 | 0.963 | 0.248 | 0.118 | 0.872 | 0.188 | 0.104 | 0.656 |
| 0.3 | 0.312 | 0.108 | 0.992 | 0.257 | 0.112 | 0.952 | 0.198 | 0.110 | 0.792 |
| 0.4 | 0.306 | 0.107 | 0.998 | 0.263 | 0.111 | 0.980 | 0.204 | 0.108 | 0.867 |
| 0.5 | 0.316 | 0.099 | 1.000 | 0.268 | 0.102 | 0.992 | 0.204 | 0.105 | 0.918 |
| 0.6 | 0.320 | 0.097 | 1.000 | 0.265 | 0.099 | 0.995 | 0.203 | 0.094 | 0.947 |
| 0.7 | 0.310 | 0.095 | 1.000 | 0.261 | 0.094 | 0.998 | 0.204 | 0.089 | 0.968 |
| 0.8 | 0.310 | 0.094 | 1.000 | 0.267 | 0.093 | 1.000 | 0.200 | 0.088 | 0.983 |
| 0.9 | 0.306 | 0.093 | 1.000 | 0.271 | 0.092 | 1.000 | 0.204 | 0.089 | 0.989 |
| 1 | 0.304 | 0.091 | 1.000 | 0.270 | 0.091 | 1.000 | 0.204 | 0.084 | 0.992 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.358 | 0.22 | 0.5164 | 0.282 | 0.1912 | 0.3736 | 0.2272 | 0.154 | 0.2804 |
| 0.04 | 0.2628 | 0.1344 | 0.5592 | 0.202 | 0.1224 | 0.3712 | 0.1576 | 0.1032 | 0.2536 |
| 0.06 | 0.2288 | 0.1116 | 0.6384 | 0.1768 | 0.1068 | 0.4324 | 0.1328 | 0.0912 | 0.2776 |
| 0.08 | 0.2136 | 0.1004 | 0.7148 | 0.1612 | 0.092 | 0.4948 | 0.1228 | 0.0852 | 0.312 |
| 0.1 | 0.2052 | 0.0948 | 0.7752 | 0.1544 | 0.0892 | 0.5624 | 0.1208 | 0.074 | 0.362 |
| 0.2 | 0.1968 | 0.0804 | 0.9344 | 0.1508 | 0.0828 | 0.7948 | 0.1104 | 0.0672 | 0.5728 |
| 0.3 | 0.2072 | 0.0792 | 0.9848 | 0.1584 | 0.0772 | 0.906 | 0.1204 | 0.0724 | 0.7132 |
| 0.4 | 0.2072 | 0.0716 | 0.996 | 0.1576 | 0.0664 | 0.9564 | 0.1232 | 0.0676 | 0.8116 |
| 0.5 | 0.2116 | 0.068 | 0.998 | 0.1604 | 0.0632 | 0.9788 | 0.1264 | 0.0652 | 0.8716 |
| 0.6 | 0.2176 | 0.068 | 1 | 0.1632 | 0.0664 | 0.9908 | 0.1228 | 0.06 | 0.914 |
| 0.7 | 0.2136 | 0.07 | 1 | 0.1596 | 0.068 | 0.998 | 0.118 | 0.06 | 0.9404 |
| 0.8 | 0.2108 | 0.072 | 1 | 0.1632 | 0.0664 | 0.9996 | 0.1224 | 0.0584 | 0.9596 |
| 0.9 | 0.2116 | 0.0684 | 1 | 0.1608 | 0.0688 | 0.9996 | 0.1248 | 0.0564 | 0.9724 |
| 1 | 0.2064 | 0.0728 | 1 | 0.162 | 0.0672 | 1 | 0.1212 | 0.0588 | 0.9856 |

Table 1.34: The Finite Sample Size of the $ExpW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP A: $(\theta, \rho, \varphi) = (0.5, 0.0, 0.0)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.368 | 0.198 | 0.712 | 0.178 | 0.135 | 0.348 | 0.108 | 0.104 | 0.179 |
| 0.04 | 0.316 | 0.201 | 0.839 | 0.160 | 0.113 | 0.494 | 0.114 | 0.108 | 0.248 |
| 0.06 | 0.301 | 0.218 | 0.900 | 0.150 | 0.113 | 0.600 | 0.111 | 0.094 | 0.314 |
| 0.08 | 0.303 | 0.221 | 0.938 | 0.144 | 0.114 | 0.694 | 0.111 | 0.090 | 0.386 |
| 0.1 | 0.291 | 0.217 | 0.956 | 0.136 | 0.113 | 0.760 | 0.104 | 0.080 | 0.454 |
| 0.2 | 0.271 | 0.154 | 0.994 | 0.137 | 0.083 | 0.915 | 0.090 | 0.058 | 0.690 |
| 0.3 | 0.271 | 0.084 | 0.999 | 0.134 | 0.051 | 0.967 | 0.089 | 0.043 | 0.817 |
| 0.4 | 0.266 | 0.048 | 1.000 | 0.129 | 0.031 | 0.986 | 0.086 | 0.028 | 0.889 |
| 0.5 | 0.256 | 0.036 | 1.000 | 0.126 | 0.023 | 0.994 | 0.086 | 0.019 | 0.929 |
| 0.6 | 0.253 | 0.026 | 1.000 | 0.124 | 0.017 | 0.998 | 0.084 | 0.017 | 0.954 |
| 0.7 | 0.254 | 0.024 | 1.000 | 0.124 | 0.016 | 1.000 | 0.082 | 0.015 | 0.976 |
| 0.8 | 0.246 | 0.020 | 1.000 | 0.125 | 0.014 | 1.000 | 0.082 | 0.012 | 0.987 |
| 0.9 | 0.250 | 0.015 | 1.000 | 0.126 | 0.010 | 1.000 | 0.084 | 0.010 | 0.993 |
| 1 | 0.254 | 0.012 | 1.000 | 0.124 | 0.007 | 1.000 | 0.082 | 0.009 | 0.996 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.094 | 0.086 | 0.404 | 0.070 | 0.072 | 0.196 | 0.062 | 0.064 | 0.115 |
| 0.04 | 0.092 | 0.082 | 0.657 | 0.070 | 0.064 | 0.345 | 0.060 | 0.058 | 0.177 |
| 0.06 | 0.090 | 0.083 | 0.774 | 0.072 | 0.064 | 0.470 | 0.058 | 0.057 | 0.247 |
| 0.08 | 0.090 | 0.079 | 0.843 | 0.068 | 0.060 | 0.572 | 0.061 | 0.058 | 0.308 |
| 0.1 | 0.086 | 0.078 | 0.888 | 0.064 | 0.058 | 0.655 | 0.060 | 0.055 | 0.367 |
| 0.2 | 0.082 | 0.063 | 0.978 | 0.060 | 0.053 | 0.866 | 0.055 | 0.044 | 0.632 |
| 0.3 | 0.082 | 0.056 | 0.994 | 0.060 | 0.052 | 0.946 | 0.062 | 0.052 | 0.779 |
| 0.4 | 0.081 | 0.047 | 0.999 | 0.058 | 0.041 | 0.974 | 0.057 | 0.050 | 0.860 |
| 0.5 | 0.083 | 0.046 | 1.000 | 0.058 | 0.038 | 0.987 | 0.055 | 0.042 | 0.906 |
| 0.6 | 0.081 | 0.034 | 1.000 | 0.059 | 0.034 | 0.995 | 0.050 | 0.041 | 0.941 |
| 0.7 | 0.081 | 0.026 | 1.000 | 0.062 | 0.028 | 1.000 | 0.056 | 0.035 | 0.960 |
| 0.8 | 0.079 | 0.025 | 1.000 | 0.059 | 0.027 | 1.000 | 0.055 | 0.031 | 0.975 |
| 0.9 | 0.080 | 0.026 | 1.000 | 0.059 | 0.025 | 1.000 | 0.051 | 0.030 | 0.987 |
| 1 | 0.082 | 0.026 | 1.000 | 0.062 | 0.024 | 1.000 | 0.055 | 0.028 | 0.992 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.0776 | 0.0776 | 0.3416 | 0.0536 | 0.0488 | 0.1668 | 0.0592 | 0.0572 | 0.1128 |
| 0.04 | 0.076 | 0.072 | 0.5936 | 0.0544 | 0.0552 | 0.308 | 0.06 | 0.0604 | 0.174 |
| 0.06 | 0.0784 | 0.0716 | 0.736 | 0.0528 | 0.0508 | 0.4544 | 0.0592 | 0.056 | 0.2368 |
| 0.08 | 0.076 | 0.0592 | 0.8136 | 0.0464 | 0.0484 | 0.5532 | 0.0592 | 0.0564 | 0.298 |
| 0.1 | 0.0692 | 0.0664 | 0.8692 | 0.0472 | 0.0536 | 0.6348 | 0.056 | 0.052 | 0.3624 |
| 0.2 | 0.0696 | 0.0584 | 0.9728 | 0.0496 | 0.0556 | 0.8584 | 0.0488 | 0.0472 | 0.62 |
| 0.3 | 0.0692 | 0.0472 | 0.9928 | 0.0492 | 0.0476 | 0.93 | 0.0488 | 0.0436 | 0.7656 |
| 0.4 | 0.07 | 0.0448 | 0.9984 | 0.0448 | 0.0464 | 0.9684 | 0.0504 | 0.05 | 0.8516 |
| 0.5 | 0.0712 | 0.0468 | 0.9996 | 0.0512 | 0.0452 | 0.9832 | 0.0544 | 0.0548 | 0.8992 |
| 0.6 | 0.072 | 0.0456 | 1 | 0.0468 | 0.0484 | 0.992 | 0.0464 | 0.05 | 0.9288 |
| 0.7 | 0.0724 | 0.0472 | 1 | 0.0464 | 0.0492 | 0.9964 | 0.0504 | 0.0468 | 0.956 |
| 0.8 | 0.0668 | 0.0436 | 1 | 0.0544 | 0.0484 | 0.9996 | 0.0476 | 0.0468 | 0.9768 |
| 0.9 | 0.07 | 0.04 | 1 | 0.0508 | 0.0456 | 0.9996 | 0.0484 | 0.0456 | 0.9872 |
| 1 | 0.0708 | 0.04 | 1 | 0.052 | 0.044 | 1 | 0.048 | 0.0424 | 0.9928 |

Table 1.35: The Finite Sample Size of the $ExpW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP D: $(\theta, \rho, \varphi) = (0.8, 0.5, 0.5)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | AP | Fixed-*b* | | AP | Fixed-*b* | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.708 | 0.333 | 0.954 | 0.646 | 0.452 | 0.818 | 0.518 | 0.398 | 0.627 |
| 0.04 | 0.517 | 0.183 | 0.960 | 0.467 | 0.226 | 0.817 | 0.360 | 0.243 | 0.568 |
| 0.06 | 0.435 | 0.172 | 0.975 | 0.398 | 0.182 | 0.861 | 0.310 | 0.195 | 0.600 |
| 0.08 | 0.396 | 0.164 | 0.982 | 0.353 | 0.164 | 0.896 | 0.292 | 0.163 | 0.651 |
| 0.1 | 0.372 | 0.165 | 0.986 | 0.332 | 0.159 | 0.917 | 0.275 | 0.139 | 0.689 |
| 0.2 | 0.324 | 0.116 | 0.997 | 0.300 | 0.102 | 0.973 | 0.235 | 0.079 | 0.844 |
| 0.3 | 0.315 | 0.072 | 1.000 | 0.283 | 0.062 | 0.994 | 0.225 | 0.054 | 0.916 |
| 0.4 | 0.307 | 0.044 | 1.000 | 0.280 | 0.036 | 0.998 | 0.226 | 0.037 | 0.954 |
| 0.5 | 0.309 | 0.034 | 1.000 | 0.281 | 0.026 | 1.000 | 0.219 | 0.034 | 0.970 |
| 0.6 | 0.312 | 0.024 | 1.000 | 0.281 | 0.023 | 1.000 | 0.216 | 0.028 | 0.982 |
| 0.7 | 0.315 | 0.020 | 1.000 | 0.273 | 0.020 | 1.000 | 0.220 | 0.024 | 0.990 |
| 0.8 | 0.310 | 0.021 | 1.000 | 0.272 | 0.021 | 1.000 | 0.215 | 0.020 | 0.996 |
| 0.9 | 0.308 | 0.021 | 1.000 | 0.270 | 0.020 | 1.000 | 0.218 | 0.020 | 0.999 |
| 1 | 0.313 | 0.020 | 1.000 | 0.278 | 0.021 | 1.000 | 0.214 | 0.018 | 1.000 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | AP | Fixed-*b* | | AP | Fixed-*b* | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.260 | 0.113 | 0.647 | 0.188 | 0.122 | 0.368 | 0.132 | 0.094 | 0.211 |
| 0.04 | 0.213 | 0.100 | 0.794 | 0.158 | 0.087 | 0.485 | 0.110 | 0.084 | 0.253 |
| 0.06 | 0.193 | 0.097 | 0.868 | 0.143 | 0.085 | 0.590 | 0.104 | 0.080 | 0.318 |
| 0.08 | 0.185 | 0.088 | 0.912 | 0.135 | 0.081 | 0.668 | 0.102 | 0.076 | 0.382 |
| 0.1 | 0.172 | 0.090 | 0.946 | 0.133 | 0.078 | 0.737 | 0.096 | 0.072 | 0.446 |
| 0.2 | 0.148 | 0.074 | 0.990 | 0.115 | 0.058 | 0.911 | 0.092 | 0.051 | 0.683 |
| 0.3 | 0.149 | 0.054 | 0.999 | 0.113 | 0.050 | 0.973 | 0.088 | 0.044 | 0.822 |
| 0.4 | 0.141 | 0.036 | 1.000 | 0.107 | 0.036 | 0.988 | 0.089 | 0.041 | 0.893 |
| 0.5 | 0.144 | 0.037 | 1.000 | 0.109 | 0.034 | 0.995 | 0.085 | 0.041 | 0.925 |
| 0.6 | 0.146 | 0.032 | 1.000 | 0.107 | 0.038 | 0.999 | 0.084 | 0.038 | 0.954 |
| 0.7 | 0.147 | 0.031 | 1.000 | 0.104 | 0.033 | 1.000 | 0.088 | 0.036 | 0.972 |
| 0.8 | 0.144 | 0.032 | 1.000 | 0.106 | 0.034 | 1.000 | 0.091 | 0.033 | 0.985 |
| 0.9 | 0.148 | 0.031 | 1.000 | 0.106 | 0.030 | 1.000 | 0.089 | 0.030 | 0.991 |
| 1 | 0.150 | 0.032 | 1.000 | 0.108 | 0.031 | 1.000 | 0.087 | 0.030 | 0.997 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-*b* | | AP | Fixed-*b* | | AP | Fixed-*b* | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.1652 | 0.0948 | 0.5116 | 0.1212 | 0.0824 | 0.2648 | 0.0924 | 0.0708 | 0.1576 |
| 0.04 | 0.1444 | 0.0892 | 0.7064 | 0.108 | 0.0648 | 0.3912 | 0.0848 | 0.0712 | 0.2124 |
| 0.06 | 0.1344 | 0.084 | 0.8044 | 0.0948 | 0.0564 | 0.52 | 0.0792 | 0.0712 | 0.278 |
| 0.08 | 0.1312 | 0.0796 | 0.8684 | 0.0912 | 0.0572 | 0.6192 | 0.0824 | 0.0688 | 0.3432 |
| 0.1 | 0.1224 | 0.0676 | 0.9048 | 0.0868 | 0.0568 | 0.6972 | 0.0784 | 0.06 | 0.3988 |
| 0.2 | 0.1184 | 0.0648 | 0.9792 | 0.0844 | 0.052 | 0.8804 | 0.064 | 0.0504 | 0.64 |
| 0.3 | 0.1048 | 0.0512 | 0.996 | 0.0744 | 0.058 | 0.9476 | 0.0656 | 0.0544 | 0.78 |
| 0.4 | 0.1072 | 0.0436 | 0.9988 | 0.0764 | 0.046 | 0.9788 | 0.0656 | 0.0516 | 0.8604 |
| 0.5 | 0.106 | 0.0424 | 0.9996 | 0.0768 | 0.04 | 0.9896 | 0.0656 | 0.0512 | 0.902 |
| 0.6 | 0.1056 | 0.042 | 1 | 0.0776 | 0.0392 | 0.9952 | 0.0692 | 0.0436 | 0.9336 |
| 0.7 | 0.1088 | 0.0412 | 1 | 0.0808 | 0.0392 | 0.9988 | 0.0692 | 0.0412 | 0.9636 |
| 0.8 | 0.1048 | 0.0416 | 1 | 0.0808 | 0.0404 | 0.9996 | 0.0664 | 0.0408 | 0.9808 |
| 0.9 | 0.1072 | 0.0404 | 1 | 0.082 | 0.0404 | 1 | 0.0672 | 0.0364 | 0.9888 |
| 1 | 0.1088 | 0.0404 | 1 | 0.0788 | 0.0404 | 1 | 0.0668 | 0.0396 | 0.9932 |

Table 1.36: The Finite Sample Size of the $ExpW^{(F)}$ Test with 5% Nominal Size

$H_0$ : No Structural Change in both $\beta_1$ and $\beta_2$, DGP F: $(\theta, \rho, \varphi) = (0.9, 0.9, 0.9)$

| T=100 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.958 | 0.624 | 0.999 | 0.975 | 0.909 | 0.994 | 0.951 | 0.902 | 0.971 |
| 0.04 | 0.719 | 0.212 | 0.995 | 0.850 | 0.514 | 0.981 | 0.826 | 0.689 | 0.925 |
| 0.06 | 0.544 | 0.143 | 0.993 | 0.721 | 0.323 | 0.977 | 0.729 | 0.522 | 0.910 |
| 0.08 | 0.445 | 0.106 | 0.992 | 0.624 | 0.238 | 0.977 | 0.665 | 0.402 | 0.908 |
| 0.1 | 0.370 | 0.095 | 0.993 | 0.558 | 0.191 | 0.981 | 0.617 | 0.327 | 0.918 |
| 0.2 | 0.282 | 0.074 | 0.996 | 0.465 | 0.123 | 0.992 | 0.529 | 0.164 | 0.956 |
| 0.3 | 0.296 | 0.060 | 0.998 | 0.458 | 0.081 | 0.997 | 0.519 | 0.105 | 0.980 |
| 0.4 | 0.307 | 0.045 | 1.000 | 0.456 | 0.057 | 0.998 | 0.505 | 0.082 | 0.991 |
| 0.5 | 0.313 | 0.036 | 1.000 | 0.456 | 0.046 | 1.000 | 0.502 | 0.064 | 0.996 |
| 0.6 | 0.304 | 0.034 | 1.000 | 0.456 | 0.038 | 1.000 | 0.507 | 0.052 | 0.998 |
| 0.7 | 0.300 | 0.033 | 1.000 | 0.451 | 0.036 | 1.000 | 0.507 | 0.042 | 0.999 |
| 0.8 | 0.297 | 0.030 | 1.000 | 0.449 | 0.034 | 1.000 | 0.500 | 0.038 | 0.999 |
| 0.9 | 0.296 | 0.026 | 1.000 | 0.451 | 0.033 | 1.000 | 0.502 | 0.035 | 1.000 |
| 1 | 0.301 | 0.026 | 1.000 | 0.452 | 0.030 | 1.000 | 0.498 | 0.032 | 1.000 |

| T=500 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.634 | 0.204 | 0.942 | 0.635 | 0.392 | 0.823 | 0.517 | 0.372 | 0.641 |
| 0.04 | 0.375 | 0.096 | 0.938 | 0.411 | 0.154 | 0.785 | 0.328 | 0.206 | 0.542 |
| 0.06 | 0.292 | 0.076 | 0.952 | 0.330 | 0.112 | 0.817 | 0.269 | 0.150 | 0.554 |
| 0.08 | 0.251 | 0.070 | 0.965 | 0.280 | 0.093 | 0.850 | 0.242 | 0.121 | 0.596 |
| 0.1 | 0.220 | 0.060 | 0.970 | 0.252 | 0.085 | 0.878 | 0.218 | 0.102 | 0.637 |
| 0.2 | 0.188 | 0.054 | 0.992 | 0.218 | 0.062 | 0.956 | 0.188 | 0.068 | 0.807 |
| 0.3 | 0.181 | 0.050 | 0.999 | 0.206 | 0.056 | 0.986 | 0.190 | 0.062 | 0.885 |
| 0.4 | 0.181 | 0.040 | 1.000 | 0.202 | 0.050 | 0.995 | 0.188 | 0.059 | 0.933 |
| 0.5 | 0.177 | 0.041 | 1.000 | 0.205 | 0.043 | 0.999 | 0.183 | 0.050 | 0.961 |
| 0.6 | 0.182 | 0.039 | 1.000 | 0.196 | 0.042 | 1.000 | 0.182 | 0.048 | 0.976 |
| 0.7 | 0.178 | 0.042 | 1.000 | 0.200 | 0.044 | 1.000 | 0.179 | 0.044 | 0.988 |
| 0.8 | 0.172 | 0.038 | 1.000 | 0.199 | 0.042 | 1.000 | 0.180 | 0.043 | 0.993 |
| 0.9 | 0.174 | 0.038 | 1.000 | 0.201 | 0.040 | 1.000 | 0.175 | 0.043 | 0.996 |
| 1 | 0.174 | 0.038 | 1.000 | 0.202 | 0.039 | 1.000 | 0.177 | 0.044 | 0.998 |

| T=1000 | $\epsilon = 0.05$ | | | $\epsilon = 0.1$ | | | $\epsilon = 0.2$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Fixed-$b$ | | AP | Fixed-$b$ | | AP | Fixed-$b$ | | AP |
| kernel | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett | Bartlett | QS | Bartlett |
| $b = 0.02$ | 0.426 | 0.1272 | 0.8092 | 0.3728 | 0.1908 | 0.5816 | 0.2912 | 0.1968 | 0.3884 |
| 0.04 | 0.2788 | 0.0728 | 0.8676 | 0.236 | 0.0948 | 0.626 | 0.1944 | 0.1196 | 0.3728 |
| 0.06 | 0.2228 | 0.0684 | 0.9092 | 0.1928 | 0.0704 | 0.6932 | 0.1612 | 0.0968 | 0.4176 |
| 0.08 | 0.198 | 0.0568 | 0.9348 | 0.1712 | 0.0588 | 0.7588 | 0.1512 | 0.088 | 0.486 |
| 0.1 | 0.1764 | 0.052 | 0.9592 | 0.1572 | 0.0544 | 0.8084 | 0.1436 | 0.074 | 0.536 |
| 0.2 | 0.1528 | 0.046 | 0.9916 | 0.1364 | 0.0528 | 0.9288 | 0.1248 | 0.0476 | 0.7312 |
| 0.3 | 0.1488 | 0.0444 | 0.9972 | 0.1368 | 0.0448 | 0.9696 | 0.1212 | 0.0452 | 0.8356 |
| 0.4 | 0.1472 | 0.0424 | 0.9996 | 0.1312 | 0.0352 | 0.9884 | 0.1236 | 0.036 | 0.8956 |
| 0.5 | 0.1532 | 0.04 | 1 | 0.1292 | 0.0352 | 0.9956 | 0.1248 | 0.036 | 0.9328 |
| 0.6 | 0.1564 | 0.0396 | 1 | 0.1356 | 0.0388 | 0.998 | 0.1264 | 0.038 | 0.9576 |
| 0.7 | 0.1536 | 0.0408 | 1 | 0.134 | 0.0392 | 0.9996 | 0.1196 | 0.0376 | 0.9768 |
| 0.8 | 0.1416 | 0.042 | 1 | 0.1384 | 0.0412 | 0.9996 | 0.1184 | 0.0388 | 0.99 |
| 0.9 | 0.1456 | 0.042 | 1 | 0.1308 | 0.0424 | 1 | 0.1208 | 0.0388 | 0.9944 |
| 1 | 0.1448 | 0.0444 | 1 | 0.1292 | 0.0408 | 1 | 0.1248 | 0.04 | 0.9972 |

APPENDIX

# Appendix for Chapter 1

The following expression is a general representation of the HAC estimators:

$$\widehat{\Omega} = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right) \widehat{v}_t \widehat{v}_s'.$$

This representation can be rewritten in terms of the partial sum processes $\widehat{S}_t = \sum_{j=1}^{t} \widehat{v}_j$ following Kiefer and Vogelsang (2005) and Hashimzade and Vogelsang (2008) as follows. Let $M = bT$. Then for the kernels in Class 1,

$$\widehat{\Omega} = T^{-2} \sum_{t=1}^{T-1} \sum_{s=1}^{T-1} T^{-1/2} \widehat{S}_t \left(T^2 \Delta_{t,s}^2\right) T^{-1/2} \widehat{S}_s', \tag{1.40}$$

where

$$\Delta_{t,s}^2 \equiv \left(K_{t,s} - K_{t,s+1}\right) - \left(K_{t+1,s} - K_{t+1,s+1}\right),$$

$$K_{t,s} = K\left(\frac{|t-s|}{bT}\right).$$

For the Class 2 kernel (Bartlett),

$$\widehat{\Omega} = \frac{2}{bT} \sum_{t=1}^{T-1} \left(T^{-1} \widehat{S}_t \widehat{S}_t'\right) - \frac{1}{bT} \sum_{t=1}^{T-M-1} \left(T^{-1} \widehat{S}_{t+bT} \widehat{S}_t' + T^{-1} \widehat{S}_t \widehat{S}_{t+bT}'\right). \tag{1.41}$$

For the kernels in Class 3,

$$\widehat{\Omega} = T^{-2} \sum\sum_{|t-s|<bT} T^{-1} \widehat{S}_t \left(T^2 \Delta_{t,s}^2\right) \widehat{S}_s' \tag{1.42}$$

$$+\frac{1}{bT} \sum_{s=1}^{T-bT} T^{-1/2} \widehat{S}_s T^{-1/2} \widehat{S}_{s+bT}' \left(\frac{K(1) - K(1-\frac{1}{bT})}{\frac{1}{bT}}\right)$$

$$-\frac{1}{bT} \sum_{s=1}^{T-bT} T^{-1/2} \widehat{S}_s T^{-1/2} \widehat{S}_{s+bT}' \left(\frac{K(-1+\frac{1}{bT}) - K(-1)}{\frac{1}{bT}}\right).$$

**Proof of Proposition 1:** From equation (1.11) and (1.12) the limit of the $\widehat{\mathbf{fi}}$ follows immediately by Assumption 1 and 2. Plugging the limits of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ in equation (1.13) yields, for $r \leq \lambda$,

$$T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \begin{pmatrix} \left(W_p(r) - \frac{r}{\lambda}W_p(\lambda)\right) \\ \mathbf{0} \end{pmatrix},$$

and for $r > \lambda$,

$$T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \left(W_p(r) - W_p(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W_p(1) - W_p(\lambda)\right)\right) \end{pmatrix}.$$

Thus one can rewrite this result by using indicator functions as

$$T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} \begin{pmatrix} F_p^{(1)}(r,\lambda) \\ F_p^{(2)}(r,\lambda) \end{pmatrix} \equiv \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \Lambda \end{pmatrix} F_p(r,\lambda),$$

where

$$F_p^{(1)}(r,\lambda) = \left(W_p(r) - \frac{r}{\lambda}W_p(\lambda)\right) \cdot \mathbf{1}(r \leq \lambda),$$

and

$$F_p^{(2)}(r,\lambda) = \left(W_p(r) - W_p(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W_p(1) - W_p(\lambda)\right)\right) \cdot \mathbf{1}(r > \lambda).$$

**Proof of Proposition 2**: From (1.18),

$$\begin{aligned}
\widehat{\Omega}^{(F)} &= \begin{pmatrix} \lambda\widehat{\Sigma}^{(1)} & T^{-1}\sum_{t=1}^{\lambda T}\sum_{s=\lambda T+1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(1)}\widehat{v}_s^{(2)\prime} \\ T^{-1}\sum_{t=\lambda T+1}^{T}\sum_{s=1}^{\lambda T} K\left(\frac{|t-s|}{M}\right)\widehat{v}_t^{(2)}\widehat{v}_s^{(1)\prime} & (1-\lambda)\widehat{\Sigma}^{(2)} \end{pmatrix} \\
&= \begin{pmatrix} \lambda\widehat{\Sigma}^{(1)} & T^{-1}\sum_{t=1}^{\lambda T}\sum_{s=\lambda T+1}^{T} K\left(\frac{|t-s|}{M}\right)x_t\widehat{u}_t\widehat{u}_s x_s' \\ T^{-1}\sum_{t=\lambda T+1}^{T}\sum_{s=1}^{\lambda T} K\left(\frac{|t-s|}{M}\right)x_t\widehat{u}_t\widehat{u}_s x_s' & (1-\lambda)\widehat{\Sigma}^{(2)} \end{pmatrix}.
\end{aligned}$$

The diagonal blocks converge respectively to $\lambda\Sigma$ and $(1-\lambda)\Sigma$ in probability under traditional asymptotics. One needs to show that the off-diagonal blocks converge to zero.

First, for the Bartlett $K\left(\frac{|t-s|}{M}\right)=0$ for $|t-s|\geq M$. Therefore, the upper right off-diagonal block can be rewritten as

$$\widehat{\Omega}_{1,2}^{(F)} \equiv \frac{1}{T}\sum_{t=1}^{\lambda T}\sum_{s=\lambda T+1}^{T} K_{t,s}\widehat{v}_t^{(1)}\widehat{v}_s^{(2)\prime} = \frac{1}{T}\sum_{t=1}^{\lambda T}\widehat{v}_t^{(1)}\sum_{s=\lambda T+1}^{T} K_{t,s}\widehat{v}_s^{(2)\prime},$$

where $K_{t,s}=K\left(\frac{|t-s|}{M}\right)$. Set $a_t=\widehat{v}_t^{(1)}$ and $b_t=\sum_{s=\lambda T+1}^{T}K_{t,s}\widehat{v}_s^{(2)\prime}$ and apply the partial summation formula

$$\sum_{t=1}^{\lambda T} a_t b_t = \sum_{t=1}^{\lambda T-1}\left[\left(\sum_{s=1}^{t} a_s\right)(b_t-b_{t+1})\right]+\left(\sum_{s=1}^{\lambda T} a_s\right) b_{\lambda T}$$

to get

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{\lambda T}\widehat{v}_t^{(1)}\sum_{s=\lambda T+1}^{T} K_{t,s}\widehat{v}_s^{(2)\prime} &= \frac{1}{T}\sum_{t=1}^{\lambda T-1}\widehat{S}_t^{(1)}\left(\sum_{s=\lambda T+1}^{T}K_{t,s}\widehat{v}_s^{(2)\prime}-\sum_{s=\lambda T+1}^{T}K_{t+1,s}\widehat{v}_s^{(2)\prime}\right) \qquad (1.43)\\ &\quad+\widehat{S}_{\lambda T}^{(1)}\sum_{s=\lambda T+1}^{T}K_{\lambda T,s}\widehat{v}_s^{(2)\prime}\\ &= \frac{1}{T}\sum_{t=1}^{\lambda T-1}\widehat{S}_t^{(1)}\left(\sum_{s=\lambda T+1}^{T}K_{t,s}\widehat{v}_s^{(2)\prime}-\sum_{s=\lambda T+1}^{T}K_{t+1,s}\widehat{v}_s^{(2)\prime}\right),\end{aligned}$$

where $\widehat{S}_t^{(1)}=\sum_{j=1}^{t}\widehat{v}_j^{(1)}$ for $t\leq\lambda T$ and $\widehat{S}_t^{(2)}=\sum_{j=\lambda T+1}^{t}\widehat{v}_j^{(1)}$ for $t\geq\lambda T+1$. The above lines used $\widehat{S}_{\lambda T}^{(1)}=0$ and $\widehat{S}_T^{(2)}=0$.

One can rewrite each component of the summands of this equation:

$$\begin{aligned}\sum_{s=\lambda T+1}^{T}K_{t,s}\widehat{v}_s^{(2)\prime} &= \sum_{s=\lambda T+1}^{T-1}(K_{t,s}-K_{t,s+1})\,\widehat{S}_s^{(2)\prime}+K_{t,T}\widehat{S}_T^{(2)\prime}-K_{t,\lambda T+1}\widehat{S}_{\lambda T}^{(2)\prime}\\ &= \sum_{s=\lambda T+1}^{T-1}(K_{t,s}-K_{t,s+1})\,\widehat{S}_s^{(2)\prime},\\ \text{and}\quad \sum_{s=\lambda T+1}^{T}K_{t+1,s}\widehat{v}_s^{(2)\prime} &= \sum_{s=\lambda T+1}^{T-1}(K_{t+1,s}-K_{t+1,s+1})\,\widehat{S}_s^{(2)\prime}+K_{t+1,T}\widehat{S}_T^{(2)\prime}-K_{t+1,\lambda T+1}\widehat{S}_{\lambda T}^{(2)\prime}\\ &= \sum_{s=\lambda T+1}^{T-1}(K_{t+1,s}-K_{t+1,s+1})\,\widehat{S}_s^{(2)\prime}\end{aligned}$$

Plugging these representations into (1.43) gives

$$\widehat{\Omega}_{1,2}^{(F)} = \frac{1}{T} \sum_{t=1}^{\lambda T-1} \sum_{s=\lambda T+1}^{T-1} \widehat{S}_t^{(1)} \left( K_{t,s} - K_{t,s+1} - K_{t+1,s} + K_{t+1,s+1} \right) \widehat{S}_s^{(2)\prime}.$$

From Hashimzade and Vogelsang (2008) for the Bartlett kernel, $K_{t,s} - K_{t,s+1} - K_{t+1,s} + K_{t+1,s+1} = -\frac{1}{M}$ for $\{(t,s) : M + t = s\}$ and zero otherwise. This yields

$$\widehat{\Omega}_{1,2}^{(F)} = \frac{1}{T} \sum_{t=\lambda T-M+1}^{\lambda T-1} \left( \frac{-1}{M} \right) \widehat{S}_t^{(1)} \widehat{S}_{t+M}^{(2)\prime}.$$

Taking the matrix-maximum norm yields

$$\begin{aligned} \left\| \widehat{\Omega}_{1,2}^{(F)} \right\| &\leq \frac{1}{TM} \sum_{t=\lambda T-M+1}^{\lambda T-1} \left\| T^{-1/2} \widehat{S}_t^{(1)} \right\| \cdot \left\| T^{-1/2} \widehat{S}_{t+M}^{(2)\prime} \right\| \\ &\leq \frac{M-1}{TM} \max_{\lambda T-M+1 \leq t \leq \lambda T-1} \left\| T^{-1/2} \widehat{S}_t^{(1)} \right\| \cdot \max_{\lambda T+1 \leq s \leq \lambda T-1+M} \left\| T^{-1/2} \widehat{S}_s^{(2)\prime} \right\| \\ &= \frac{1}{T} O_p(1) = o_p(1). \end{aligned}$$

Here we used

$$\max_{\lambda T-M+1 \leq t \leq \lambda T-1} \left\| T^{-1/2} \widehat{S}_t^{(1)} \right\| = O_p(1),$$

and

$$\max_{\lambda T+1 \leq s \leq \lambda T-1+M} \left\| T^{-1/2} \widehat{S}_s^{(2)\prime} \right\| = O_p(1),$$

which is true given the result in Proposition 1 and the traditional assumption that $M/T \to 0$. The same argument applies to the left lower off-diagonal block of $\widehat{\Omega}^{(F)}$ yielding the desired result.

**Proof of Lemma 1**: Plugging the limit of the partial sum process in Proposition 1 into the HAC estimators in (1.40), (1.41), or (1.42) the desired result follows from direct application of the continuous mapping theorem to obtain the desired result in (1.19).

**Proof of Lemma 2**: The proof is for the Bartlett kernel. Recall from (1.16) that

$$\widehat{\Sigma}^{(1)} = \frac{1}{\lambda T} \sum_{t=1}^{[\lambda T]} \sum_{s=1}^{[\lambda T]} K\left(\frac{|t-s|}{M_1}\right) \widehat{v}_t^{(1)} \widehat{v}_s^{(1)\prime}.$$

With the Bartlett kernel, one can rewrite this as:

$$\widehat{\Sigma}^{(1)} = \frac{2}{b_1} \frac{1}{\lambda T} \sum_{t=1}^{[\lambda T]-1} \left( (\lambda T)^{-1/2} \widehat{S}_t^{(1)} (\lambda T)^{-1/2} \widehat{S}_t^{(1)\prime} \right)$$

$$-\frac{1}{b_1} \frac{1}{\lambda T} \sum_{t=1}^{[(1-b_1)\lambda T]-1} \left\{ \left( (\lambda T)^{-1/2} \widehat{S}_{t+b_1\lambda T}^{(1)} (\lambda T)^{-1/2} \widehat{S}_t^{(1)\prime} \right) + \left( (\lambda T)^{-1/2} \widehat{S}_t^{(1)} (\lambda T)^{-1/2} \widehat{S}_{t+b_1\lambda T}^{(1)\prime} \right) \right\},$$

with $b_1 = \frac{M_1}{\lambda T}$ and $\widehat{S}_t^{(1)} = \sum_{j=1}^{t} \widehat{v}_j^{(1)}$. Apply the continuous mapping theorem using the result on the limit of the partial sum process in Proposition 1 to obtain

$$\widehat{\Sigma}^{(1)} \Rightarrow \frac{2}{b_1 \lambda} \int_0^1 \lambda^{-1/2} \Lambda F_p^{(1)}(r,\lambda) F_p^{(1)}(r,\lambda)' \lambda^{-1/2} \Lambda' dr$$

$$-\frac{1}{b_1\lambda} \int_0^{(1-b_1)\lambda} \lambda^{-1/2} \Lambda F_p^{(1)}(r + b_1\lambda, \lambda) F_p^{(1)}(r,\lambda)' \lambda^{-1/2} \Lambda' dr$$

$$-\frac{1}{b_1\lambda} \int_0^{(1-b_1)\lambda} \lambda^{-1/2} \Lambda F_p^{(1)}(r,\lambda) F_p^{(1)}(r + b_1\lambda, \lambda)' \lambda^{-1/2} \Lambda' dr$$

$$= \Lambda \left( \frac{\mathbf{P}\left(b_1\lambda, F_p^{(1)}(r,\lambda)\right)}{\lambda} \right) \Lambda'$$

by recalling (1.4). One can easily get the limit of $\widehat{\Sigma}^{(2)}$ in the same way and this completes the proof.

**Proof of Theorem 1:** Recall that

$$Wald^{(F)} = T\left(R\widehat{\boldsymbol{\beta}}\right)' \left(R\widehat{\mathbf{Q}}_\lambda^{-1} \widehat{\Omega}^{(F)} \widehat{\mathbf{Q}}_\lambda^{-1} R'\right)^{-1} \left(R\widehat{\boldsymbol{\beta}}\right).$$

Using $R = (R_1, \ -R_1)$ it follows that

$$RT^{1/2}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \overset{H_0}{=} R_1\left(T^{1/2}\left(\widehat{\beta}_1 - \beta_1\right) - T^{1/2}\left(\widehat{\beta}_2 - \beta_2\right)\right) \overset{d}{\to}$$

$$R_1 Q^{-1}\Lambda\left(\frac{1}{\lambda}W_p(\lambda) - \frac{1}{1-\lambda}\left(W_p(1) - W_p(\lambda)\right)\right).$$

Using Assumption 1 and Lemma 1,

$$R\widehat{\mathbf{Q}}_\lambda^{-1}\widehat{\Omega}^{(F)}\widehat{\mathbf{Q}}_\lambda^{-1}R' \Rightarrow \left(\frac{1}{\lambda}R_1Q^{-1}\Lambda,\ \frac{-1}{1-\lambda}R_1Q^{-1}\Lambda\right) \times \mathbf{P}\left(b, F_p(r,\lambda)\right)$$

$$\times\left(\frac{1}{\lambda}R_1Q^{-1}\Lambda,\ \frac{-1}{1-\lambda}R_1Q^{-1}\Lambda\right)'.$$

By writing out $\mathbf{P}\left(b, F_p(r,\lambda)\right)$ using $F_p(r,\lambda) = \left(F_p^{(1)}(r,\lambda)',\ F_p^{(2)}(r,\lambda)'\right)'$, one can obtain the following expression for above limit after some algebra:

$$R_1Q^{-1}\Lambda\mathbf{P}\left(b, \frac{1}{\lambda}F_p^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_p^{(2)}(r,\lambda)\right)\Lambda'Q^{-1}R_1'.$$

Now apply the transformation

$$R_1Q^{-1}\Lambda W_p(r) \overset{d}{=} AW_l(r),$$

with

$$R_1Q^{-1}\Lambda\Lambda'Q^{-1}R_1' = AA',$$

and conclude

$$R\widehat{\mathbf{Q}}_\lambda^{-1}\widehat{\Omega}^{(F)}\widehat{\mathbf{Q}}_\lambda^{-1}R' \Rightarrow A\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)A',$$

giving the needed result:

$$Wald^{(F)} \Rightarrow \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right)'$$

$$\times \left(\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\right)^{-1} \times \left(\frac{1}{\lambda}W_l(\lambda) - \frac{1}{1-\lambda}\left(W_l(1) - W_l(\lambda)\right)\right).$$

**Proof of Corollary 2:** Consider the Bartlerr kernel case. Other cases can be proved in the same way. Look at $\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)$ in (1.24). Rewrite it using the definitions of $F_l^{(1)}(r,\lambda)$ and $F_l^{(2)}(r,\lambda)$ in Lemma 1.

$$\mathbf{P}\left(b, \frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)$$

$$= \frac{2}{b}\int_0^1 \left(\frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\left(\frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)' dr$$

$$-\frac{1}{b}\int_0^{1-b} \left(\frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)\left(\frac{1}{\lambda}F_l^{(1)}(r+b,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r+b,\lambda)\right)' dr$$

$$-\frac{1}{b}\int_0^{1-b} \left(\frac{1}{\lambda}F_l^{(1)}(r+b,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r+b,\lambda)\right)\left(\frac{1}{\lambda}F_l^{(1)}(r,\lambda) - \frac{1}{1-\lambda}F_l^{(2)}(r,\lambda)\right)' dr.$$

Rewrite the above expression by noting that $F_l^{(1)}(r,\lambda) \times F_l^{(2)}(r,\lambda) = 0$ as

$$\frac{2}{b}\int_0^1 \frac{1}{\lambda^2}F_l^{(1)}(r,\lambda)\, F_l^{(1)}(r,\lambda)' dr$$

$$-\frac{1}{b}\int_0^{1-b} \left(\frac{1}{\lambda^2}F_l^{(1)}(r,\lambda)\, F_l^{(1)}(r+b,\lambda)' + \frac{1}{\lambda^2}F_l^{(1)}(r+b,\lambda)\, F_l^{(1)}(r,\lambda)'\right) dr$$

$$+\frac{2}{b}\int_0^1 \frac{1}{(1-\lambda)^2}F_l^{(2)}(r,\lambda)\, F_l^{(2)}(r,\lambda)' dr - \frac{1}{b}\int_0^{1-b} \left(\frac{1}{(1-\lambda)^2}F_l^{(2)}(r,\lambda)\, F_l^{(2)}(r+b,\lambda)'\right.$$

$$\left.+\frac{1}{(1-\lambda)^2}F_l^{(2)}(r+b,\lambda)\, F_l^{(2)}(r,\lambda)' dr\right) dr + \frac{1}{b}\int_0^{1-b} \frac{1}{\lambda(1-\lambda)}\left(F_l^{(1)}(r,\lambda)\, F_l^{(2)}(r+b,\lambda)'\right.$$

$$\left.+\, F_l^{(2)}(r,\lambda)\, F_l^{(1)}(r+b,\lambda)' + F_l^{(1)}(r+b,\lambda)\, F_l^{(2)}(r,\lambda)' + F_l^{(2)}(r+b,\lambda)\, F_l^{(1)}(r,\lambda)'\right) dr.$$

By reminding that $\lambda$ is given fixed and noticing $\left\{F_l^{(1)}(r,\lambda)\right\}_{0\leq r\leq\lambda}$ and $\left\{F_l^{(2)}(r,\lambda)\right\}_{\lambda\leq r\leq 1}$ are independent, one can still preserve the distribution of the above expression under the following transformations

$$\begin{aligned}
&\sqrt{\lambda}W_l^{*}\left(\frac{r}{\lambda}\right)\overset{d}{=}W_l(r)\text{ for }W\left(\cdot\right)\text{ in }F_l^{(1)}(r,\lambda),\\
&\sqrt{1-\lambda}W_l^{**}\left(\frac{r-\lambda}{1-\lambda}\right)\overset{d}{=}W_l(r)-W_l(\lambda)\text{ for }W\left(\cdot\right)\text{ in }F_l^{(2)}(r,\lambda).
\end{aligned}$$

Then apply change of variables: $\frac{r}{\lambda}=u$ for the first two integrals and $\frac{r-\lambda}{1-\lambda}=v$ for the next two integrals. Finally, notice the term outside the inverse in equation (1.24) is independent of the term inside the inverse. So one can apply the following separate transformation to the outside term:

$$\frac{1}{\lambda}W_l(\lambda)-\frac{1}{1-\lambda}\left(W_l(1)-W_l(\lambda)\right)\overset{d}{=}\frac{1}{\sqrt{\lambda(1-\lambda)}}W_l(1),$$

which yields the desired result.

**Proof of Theorem 2:** Using Lemma 2 and the transformation shown in the proof of Theorem 1, the result in Theorem 2 immediately follows.

**Proof of Corollary 3:** Proof is similar to the proof of Corollary 2.

**Proof of Proposition 3:** Frisch-Waugh Theorem gives

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}&=\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\right)^{-1}\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{y}_t\right) &&(1.44)\\
&=\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\right)^{-1}\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\boldsymbol{\beta}+\sum_{t=1}^{T}\widetilde{X}_tu_t-\sum_{t=1}^{T}\widetilde{X}_tz_t'(Z'Z)^{-1}Z'u\right)\\
&=\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\right)^{-1}\left(\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\boldsymbol{\beta}+\sum_{t=1}^{T}\widetilde{X}_tu_t\right),
\end{aligned}$$

and it immediately follows that

$$T^{1/2}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)=\left(T^{-1}\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\right)^{-1}\left(T^{-1/2}\sum_{t=1}^{T}\widetilde{X}_t u_t\right). \tag{1.45}$$

$$=\left(T^{-1}\sum_{t=1}^{T}\left(X_t-X'Z(Z'Z)^{-1}z_t\right)\left(X_t'-z_t'(Z'Z)^{-1}Z'X\right)\right)^{-1} \tag{1.46}$$

$$\times\left(T^{-1/2}\sum_{t=1}^{T}\left(X_t-X'Z(Z'Z)^{-1}z_t\right)u_t\right). \tag{1.47}$$

Under Assumption 1′ and 2′ it follows in a straightforward manner that

$$\sqrt{T}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\Rightarrow Q_{\widetilde{X}\widetilde{X}}^{-1}\begin{pmatrix}\Lambda_1 W_{p+q}(\lambda)-\lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1)\\ \Lambda_1\left(W_{p+q}(1)-W_{p+q}(\lambda)\right)-(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1)\end{pmatrix} \tag{1.48}$$

In order to derive the limit of $\sqrt{T}(R\widehat{\boldsymbol{\beta}}-r)$ following immediate results are useful:

$$Q_{XZ}\equiv p\lim\left(T^{-1}\sum_{t=1}^{T}X_t z_t'\right)=\begin{pmatrix}\lambda Q_{xZ}\\(1-\lambda)Q_{xZ}\end{pmatrix}_{2p\times q},$$

$$Q_{XX}\equiv p\lim\left(T^{-1}\sum_{t=1}^{T}X_t X_t'\right)=\begin{pmatrix}\lambda Q_{xx} & \mathbf{0}\\ \mathbf{0} & (1-\lambda)Q_{xx}\end{pmatrix}_{2p\times 2p},$$

$$Q_{\widetilde{X}\widetilde{X}}=p\lim\left(T^{-1}\sum_{t=1}^{T}\widetilde{X}_t\widetilde{X}_t'\right)=Q_{XX}-Q_{XZ}Q_{ZZ}^{-1}Q_{XZ}'.$$

Also by recalling from matrix algebra (see e.g. Schott (1997)) that

$$Q_{\widetilde{X}\widetilde{X}}^{-1}=Q_{XX}^{-1}+Q_{XX}^{-1}Q_{XZ}\left(Q_{ZZ}-Q_{XZ}'Q_{XX}^{-1}Q_{XZ}\right)^{-1}Q_{XZ}'Q_{XX}^{-1} \tag{1.49}$$

and using (1.49) one can further show that

$$Q^{-1}_{\widetilde{X}\widetilde{X}} = \begin{pmatrix} \frac{1}{\lambda} Q^{-1}_{xx} + P & P \\ P & \frac{1}{1-\lambda} Q^{-1}_{xx} + P \end{pmatrix}, \tag{1.50}$$

where

$$P = Q^{-1}_{xx} Q_{xZ} \left( Q_{ZZ} - Q'_{xZ} Q^{-1}_{xx} Q_{xZ} \right)^{-1} Q'_{xZ} Q^{-1}_{xx}.$$

Now plug (1.50) in (1.48) and conclude

$$\begin{aligned} &\sqrt{T}(R\widehat{\boldsymbol{\beta}} - \mathbf{r}) \stackrel{H_0}{=} \sqrt{T} R \left( \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \\ &\Rightarrow R_1 Q^{-1}_{xx} \Lambda_1 \left( \frac{1}{\lambda} W_{p+q}(\lambda) + \frac{1}{1-\lambda} \left( W_{p+q}(\lambda) - W_{p+q}(1) \right) \right). \end{aligned} \tag{1.51}$$

The following lemma is used in the proof of Lemma 3.

**Lemma 4.** *Let $K = Q_{xZ} Q^{-1}_{ZZ} Q'_{xZ}$. Then it holds that $Q^{-1}_{xx} K P = P - Q^{-1}_{xx} K Q^{-1}_{xx}$.*

**Proof of Lemma 4:** From equation (1.46),

$$Q_{\widetilde{X}\widetilde{X}} = \begin{pmatrix} \lambda Q_{xx} & 0 \\ 0 & (1-\lambda) Q_{xx} \end{pmatrix} - \begin{pmatrix} \lambda^2 Q_{xZ} Q^{-1}_{ZZ} Q'_{xZ} & \lambda(1-\lambda) Q_{xZ} Q^{-1}_{ZZ} Q'_{xZ} \\ \lambda(1-\lambda) Q_{xZ} Q^{-1}_{ZZ} Q'_{xZ} & (1-\lambda)^2 Q_{xZ} Q^{-1}_{ZZ} Q'_{xZ} \end{pmatrix}.$$

The desired result comes from the identity $Q_{\widetilde{X}\widetilde{X}} Q^{-1}_{\widetilde{X}\widetilde{X}} = I$ by substituting equation (1.50) for $Q^{-1}_{\widetilde{X}\widetilde{X}}$.

**Proof of Lemma 3:** First note that implicit in the proof of Proposition 3 is the result that $p\lim \widehat{Q}^{-1}_{\widetilde{X}\widetilde{X}} = Q^{-1}_{\widetilde{X}\widetilde{X}}$. For $R = (R_1, \ -R_1)$ it follows that

$$p\lim R \widehat{Q}^{-1}_{\widetilde{X}\widetilde{X}} = R_1 \left( \frac{1}{\lambda} Q^{-1}_{xx}, \ -\frac{1}{1-\lambda} Q^{-1}_{xx} \right) \tag{1.52}$$

using (1.50). The scaled partial sum process is given by

$$T^{-1/2}\widehat{S}^{\xi}_{[rT]} = T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t\widehat{u}_t$$

$$= T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t u_t - T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t\widetilde{X}_t'\sqrt{T}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) - T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t z_t'\left(\frac{Z'Z}{T}\right)^{-1}\left(T^{-1/2}Z'u\right). \tag{1.53}$$

For $0 \leq r < \lambda$, the first term in (1.53) satisfies

$$T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t u_t \Rightarrow \begin{pmatrix}\Lambda_1 W_{p+q}(r) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\\ -(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\end{pmatrix}. \tag{1.54}$$

Hence with $R = (R_1,\ -R_1)$, from (1.52) and (1.54) it follows that

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t u_t \Rightarrow R_1\left(\frac{1}{\lambda}Q_{xx}^{-1},\ -\frac{1}{1-\lambda}Q_{xx}^{-1}\right)\begin{pmatrix}\Lambda_1 W_{p+q}(r) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\\ -(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\end{pmatrix}$$

$$= \frac{1}{\lambda}R_1 Q_{xx}^{-1}\Lambda_1 W_{p+q}(r).$$

For the first part of the second term in (1.53) it follows that

$$T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t\widetilde{X}_t' \Rightarrow \begin{pmatrix} rQ_{xx} & \mathbf{0}_{p\times p}\\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p}\end{pmatrix} - \begin{pmatrix} r\lambda Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & r(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}'\\ \mathbf{0}_{p\times p} & \mathbf{0}_{p\times p}\end{pmatrix}$$

$$-\begin{pmatrix} r\lambda Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \mathbf{0}_{p\times p}\\ r(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \mathbf{0}_{p\times p}\end{pmatrix} + r\begin{pmatrix}\lambda^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}'\\ \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & (1-\lambda)^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}'\end{pmatrix}$$

$$= \begin{pmatrix} rQ_{xx} + (r\lambda^2 - 2r\lambda)K & -r(1-\lambda)^2 K\\ -r(1-\lambda)^2 K & r(1-\lambda)^2 K\end{pmatrix},$$

where $K = Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}'$. Hence with $R = (R_1, \ -R_1)$,

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t\widetilde{X}_t' \Rightarrow R_1\left(\frac{1}{\lambda}Q_{xx}^{-1}, \ -\frac{1}{1-\lambda}Q_{xx}^{-1}\right) \times \begin{pmatrix} rQ_{xx} + (r\lambda^2 - 2r\lambda)K & -r(1-\lambda)^2K \\ -r(1-\lambda)^2K & r(1-\lambda)^2K \end{pmatrix}$$

$$= rR_1\left(\frac{1}{\lambda}I - Q_{xx}^{-1}K, \ \frac{\lambda-1}{\lambda}Q_{xx}^{-1}K\right),$$

which combined with (1.48) and Lemma 4 immediately yields

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}\left(T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t\widetilde{X}_t'\right)\sqrt{T}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow$$

$$rR_1\left(\frac{1}{\lambda}I - Q_{xx}^{-1}K, \ \frac{\lambda-1}{\lambda}Q_{xx}^{-1}K\right) \times \begin{pmatrix} \frac{1}{\lambda}Q_{xx}^{-1} + P & P \\ P & \frac{1}{1-\lambda}Q_{xx}^{-1} + P \end{pmatrix}$$
$$\times \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1) \\ \Lambda_1\left(W_{p+q}(1) - W_{p+q}(\lambda)\right) - (1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1) \end{pmatrix}$$

$$= rR_1\left(\frac{1}{\lambda^2}Q_{xx}^{-1}, \ \mathbf{0}_{p\times p}\right) \times \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1) \\ \Lambda_1\left(W_{p+q}(1) - W_{p+q}(\lambda)\right) - (1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1) \end{pmatrix}$$
$$= \frac{r}{\lambda^2}R_1Q_{xx}^{-1}\Lambda_1 W_{p+q}(\lambda) - \frac{r}{\lambda}R_1Q_{xx}^{-1}Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1).$$

Finally, premultiplying the third term in (1.53) by $R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}$ gives

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t z_t'\left(\frac{Z'Z}{T}\right)^{-1}\left(T^{-1/2}Z'u\right)$$
$$= R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1}\sum_{t=1}^{[rT]}\left(X_t - X'Z(Z'Z)^{-1}z_t\right)z_t'\left(\frac{Z'Z}{T}\right)^{-1}\left(T^{-1/2}Z'u\right)$$

$$\Rightarrow R_1\left(\frac{1}{\lambda}Q_{xx}^{-1},\ -\frac{1}{1-\lambda}Q_{xx}^{-1}\right)\times\begin{pmatrix} r(1-\lambda)Q_{xZ} \\ -r(1-\lambda)Q_{xZ}\end{pmatrix} Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1)$$
$$= \frac{r}{\lambda}R_1 Q_{xx}^{-1}Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(1).$$

Combining the results for the three terms gives

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1/2}\widehat{S}_{[rT]}^{\xi} \Rightarrow R_1 Q_{xx}^{-1}\Lambda_1\frac{1}{\lambda}W_{p+q}(r) - R_1 Q_{xx}^{-1}\Lambda_1\frac{r}{\lambda^2}W_{p+q}(\lambda)$$
$$+R_1 Q_{xx}^{-1}Q_{xZ}Q_{ZZ}^{-1}\Lambda_2\frac{r}{\lambda}W_{p+q}(1) - R_1 Q_{xx}^{-1}Q_{xZ}Q_{ZZ}^{-1}\Lambda_2\frac{r}{\lambda}W_{p+q}(1)$$
$$= R_1 Q_{xx}^{-1}\Lambda_1\left(\frac{1}{\lambda}W_{p+q}(r) - \frac{r}{\lambda^2}W_{p+q}(\lambda)\right) = R_1 Q_{xx}^{-1}\Lambda_1\frac{1}{\lambda}F_{p+q}^{(1)}(r,\lambda). \tag{1.55}$$

Now consider $\lambda \le r \le 1$, for the first term in 1.53) it follows that

$$T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t u_t \Rightarrow \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r) \\ \Lambda_1\left(W_{p+q}(r) - W_{p+q}(\lambda)\right) - (1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\end{pmatrix}.$$

Hence with $R = (R_1,\ -R_1)$,

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}T^{-1/2}\sum_{t=1}^{[rT]}\widetilde{X}_t u_t \Rightarrow$$

$$R_1\left(\frac{1}{\lambda}Q_{xx}^{-1},\ -\frac{1}{1-\lambda}Q_{xx}^{-1}\right)\times\begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r) \\ \Lambda_1\left(W_{p+q}(r) - W_{p+q}(\lambda)\right) - (1-\lambda)Q_{xZ}Q_{ZZ}^{-1}\Lambda_2 W_{p+q}(r)\end{pmatrix}.$$
$$= R_1 Q_{xx}^{-1}\Lambda_1\left(\frac{1}{\lambda}W_{p+q}(\lambda) - \frac{1}{1-\lambda}\left(W_{p+q}(r) - W_{p+q}(\lambda)\right)\right).$$

For the first part of the second term in (1.53) it follows that

$$T^{-1}\sum_{t=1}^{[rT]}\widetilde{X}_t\widetilde{X}_t' \Rightarrow \begin{pmatrix} \lambda Q_{xx} & \mathbf{0}_{p\times p} \\ \mathbf{0}_{p\times p} & (r-\lambda)Q_{xx}\end{pmatrix}$$

$$-\begin{pmatrix} \lambda^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \\ \lambda(r-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & (r-\lambda)(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \end{pmatrix}$$

$$-\begin{pmatrix} \lambda^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \lambda(r-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \\ \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & (r-\lambda)(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \end{pmatrix}$$

$$+r\begin{pmatrix} \lambda^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \\ \lambda(1-\lambda)Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' & (1-\lambda)^2 Q_{xZ}Q_{ZZ}^{-1}Q_{xZ}' \end{pmatrix}$$

$$= \begin{pmatrix} \lambda Q_{xx} + (r-2)\lambda^2 K & \left[(2-r)\lambda^2 - \lambda\right] K \\ \left[(2-r)\lambda^2 - \lambda\right] K & (r-\lambda)Q_{xx} + \left((r-2)\,\lambda + r\right)(\lambda-1)\,K \end{pmatrix}.$$

Hence with $R = (R_1, \ -R_1)$,

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1} T^{-1} \sum_{t=1}^{[rT]} \widetilde{X}_t \widetilde{X}_t' \Rightarrow$$

$$R_1\left(\frac{1}{\lambda}Q_{xx}^{-1}, \ -\frac{1}{1-\lambda}Q_{xx}^{-1}\right) \times \begin{pmatrix} \lambda Q_{xx} + (r-2)\lambda^2 K & \left[(2-r)\lambda^2 - \lambda\right] K \\ \left[(2-r)\lambda^2 - \lambda\right] K & (r-\lambda)Q_{xx} + \left((r-2)\,\lambda + r\right)(\lambda-1)\,K \end{pmatrix}$$

$$= R_1\left(I + \frac{\lambda(r-1)}{1-\lambda}Q_{xx}^{-1}K, \ -\frac{r-\lambda}{1-\lambda}I + (r-1)Q_{xx}^{-1}K\right).$$

It directly follows that

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}\left(T^{-1}\sum_{t=1}^{[rT]} \widetilde{X}_t \widetilde{X}_t'\right)\sqrt{T}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow$$

$$R_1 \left( I + \frac{\lambda(r-1)}{1-\lambda} Q_{xx}^{-1} K, \ -\frac{r-\lambda}{1-\lambda} I + (r-1) Q_{xx}^{-1} K \right) \times \begin{pmatrix} \frac{1}{\lambda} Q_{xx}^{-1} + P & P \\ P & \frac{1}{1-\lambda} Q_{xx}^{-1} + P \end{pmatrix}$$

$$\times \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \\ \Lambda_1 \left( W_{p+q}(1) - W_{p+q}(\lambda) \right) - (1-\lambda) Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \end{pmatrix}$$

$$= R_1 \left( \frac{1}{\lambda} Q_{xx}^{-1}, \ -\frac{r-\lambda}{(1-\lambda)^2} Q_{xx}^{-1} \right)$$

$$\times \begin{pmatrix} \Lambda_1 W_{p+q}(\lambda) - \lambda Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \\ \Lambda_1 \left( W_{p+q}(1) - W_{p+q}(\lambda) \right) - (1-\lambda) Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1) \end{pmatrix}$$

$$= R_1 Q_{xx}^{-1} \Lambda_1 \left( \frac{1}{\lambda} W_{p+q}(\lambda) + \frac{r-\lambda}{(1-\lambda)^2} W_{p+q}(\lambda) - \frac{r-\lambda}{(1-\lambda)^2} W_{p+q}(1) \right)$$

$$-R_1 Q_{xx}^{-1} Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 \frac{1-r}{1-\lambda} W_{p+q}(1).$$

Finally, premultiplying the third term in (1.53) by $R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1}$ gives

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1} T^{-1} \sum_{t=1}^{[rT]} \widetilde{X}_t z_t' \left( \frac{Z'Z}{T} \right)^{-1} \left( T^{-1/2} Z'u \right)$$

$$= R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1} T^{-1} \sum_{t=1}^{[rT]} \left( X_t - X'Z(Z'Z)^{-1} z_t \right) z_t' \left( \frac{Z'Z}{T} \right)^{-1} \left( T^{-1/2} Z'u \right)$$

$$\Rightarrow R_1 \left( \frac{1}{\lambda} Q_{xx}^{-1}, \ -\frac{1}{1-\lambda} Q_{xx}^{-1} \right) \times \begin{pmatrix} \lambda(1-r) Q_{xZ} \\ -\lambda(1-r) Q_{xZ} \end{pmatrix} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1)$$

$$= \frac{1-r}{1-\lambda} R_1 Q_{xx}^{-1} Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 W_{p+q}(1).$$

Combining the results for the three terms gives

$$R\widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1} T^{-1/2} \widehat{S}_{[rT]}^{\xi} \Rightarrow R_1 Q_{xx}^{-1} \Lambda_1$$

$$\times \left( -\frac{1}{1-\lambda} \left( W(r) - W(\lambda) \right) - \frac{r-\lambda}{(1-\lambda)^2} W(\lambda) + \frac{r-\lambda}{(1-\lambda)^2} W(1) \right)$$

$$+ R_1 Q_{xx}^{-1} Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 \frac{1-r}{1-\lambda} W(1) - R_1 Q_{xx}^{-1} Q_{xZ} Q_{ZZ}^{-1} \Lambda_2 \cdot \frac{1-r}{1-\lambda} W(1)$$

$$= R_1 Q_{xx}^{-1} \Lambda_1 \left( -\frac{1}{1-\lambda} \left( W(r) - W(\lambda) \right) - \frac{r-\lambda}{(1-\lambda)^2} W(\lambda) + \frac{r-\lambda}{(1-\lambda)^2} W(1) \right)$$

$$= -R_1 Q_{xx}^{-1} \Lambda_1 \cdot \frac{1}{1-\lambda} F_{p+q}^{(2)}(r, \lambda). \tag{1.56}$$

By combining (1.55) and (1.56), for $r \in [0.1]$,

$$R \widehat{Q}_{\widetilde{X}\widetilde{X}}^{-1} T^{-1/2} \widehat{S}_{[Tr]}^{\xi} \Rightarrow R_1 Q_{xx}^{-1} \Lambda_1 \left( \frac{1}{\lambda} F_{p+q}^{(1)}\left(r, \lambda\right) - \frac{1}{1-\lambda} F_{p+q}^{(2)}\left(r, \lambda\right) \right).$$

**Proof of Theorem 3:** To conserve on space the proof for this Theorem is provided only for the case of the Bartlett kernel with $M = T$ (i.e. $b = 1$). However, the proof given here goes through for other kernels and different values of $b$. Note that with $b = 1$ the HAC estimator in equation (1.41) can be rewritten as

$$\widehat{\Omega}_{b=1}^{(F)} = \frac{2}{T} \sum_{t=1}^{T-1} T^{-1/2} \widehat{S}_t^{\xi} T^{-1/2} \widehat{S}_t^{\xi\prime}.$$

With this HAC estimator the term within the inverse in (1.35) is given by

$$\frac{2}{T} \sum_{t=1}^{T-1} \left\{ R \left( T^{-1} \sum_{s=1}^{T} \widetilde{X}_s \widetilde{X}_s' \right)^{-1} T^{-1/2} \widehat{S}_t^{\xi} T^{-1/2} \widehat{S}_t^{\xi\prime} \left( T^{-1} \sum_{s=1}^{T} \widetilde{X}_s \widetilde{X}_s' \right)^{-1} R' \right\}$$
$$\Rightarrow \mathbf{P} \left( 1, R_1 Q_{xx}^{-1} \Lambda_1 \left( \frac{1}{\lambda} F_{p+q}^{(1)}\left(r, \lambda\right) - \frac{1}{1-\lambda} F_{p+q}^{(2)}\left(r, \lambda\right) \right) \right)$$

where the limit is obtained directly from Lemma 3 and the continuous mapping theorem. The result for (1.35) can be obtained by using similar arguments as those used in Theorem 1 where the transformation is used: $R_1 Q_{xx}^{-1} \Lambda_1 W_{p+q}(r) \overset{d}{=} \Xi \cdot W_l(r)$, $0 \leq r \leq 1$ for a p.d. matrix $\underset{l \times l}{\Xi}$ satisfying $\Xi\Xi' = R_1 Q_{xx}^{-1} \Lambda_1 \Lambda_1' Q_{xx}^{-1} R_1'$.

# REFERENCES

REFERENCES


Andrews, D. W. K.: (1991), 'Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation'. *Econometrica* **59**, 817–854.

Andrews, D. W. K.: (1993), 'Tests for Parameter Instability and Structural Change with Unknown Change Point'. *Econometrica* **61**, 821–856.

Andrews, D. W. K. and W. Ploberger: (1994), 'Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative'. *Econometrica* **62**, 1383–1414.

Bai, J. and P. Perron: (2006), 'Multiple structural change models: a simulation analysis'. *Econometric theory and practice: Frontiers of analysis and applied research* pp. 212–237.

Bai, J. S. and P. Perron: (1998), 'Estimating and Testing Linear Models with Multiple Structural Breaks'. *Econometrica* **66**, 47–78.

Davidson, J.: (1994), *Stochastic Limit Theory*. New York: Oxford University Press.

De Jong, R. M. and J. Davidson: (2000), 'Consistency of Kernel Estimators of Heteroskedastic and Autocorrelated Covariance Matrices'. *Econometrica* **68**, 407–424.

Gonçalves, S. and T. J. Vogelsang: (2011), 'Block Bootstrap HAC Robust Tests: The Sophistication of the Naive Bootstrap'. *Econometric Theory* **27**(4), 745–791.

Hansen, B. E.: (1992), 'Consistent Covariance Matrix Estimation for Dependent Heterogenous Processes'. *Econometrica* **60**, 967–972.

Hashimzade, N. and T. J. Vogelsang: (2008), 'Fixed-$b$ Asymptotic Approximation of the Sampling Behavior of Nonparametric Spectral Density Estimators'. *Journal of Time Series Analysis* **29**, 142–162.

Jansson, M.: (2002), 'Consistent Covariance Estimation for Linear Processes'. *Econometric Theory* **18**, 1449–1459.

Jansson, M.: (2004), 'The Error Rejection Probability of Simple Autocorrelation Robust Tests'. *Econometrica* **72**, 937–946.

Kiefer, N. M. and T. J. Vogelsang: (2002a), 'Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation'. *Econometrica* **70**, 2093–2095.

Kiefer, N. M. and T. J. Vogelsang: (2002b), 'Heteroskedasticity-Autocorrelation Robust Testing Using Bandwidth Equal to Sample Size'. *Econometric Theory* **18**, 1350–1366.

Kiefer, N. M. and T. J. Vogelsang: (2005), 'A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests'. *Econometric Theory* **21**, 1130–1164.

Newey, W. K. and K. D. West: (1987), 'A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix'. *Econometrica* **55**, 703–708.

Perron, P.: (2006), 'Dealing with structural breaks'. *Palgrave handbook of econometrics* **1**, 278–352.

Phillips, P. C. B. and S. N. Durlauf: (1986), 'Multiple Regression with Integrated Processes'. *Review of Economic Studies* **53**, 473–496.

Phillips, P. C. B. and V. Solo: (1992), 'Asymptotics for Linear Processes'. *The Annals of Statistics* **20**, 971–1001.

Schott, J. R.: (1997), 'Matrix analysis for statistics'. *A Wiley InterScience Publication, New York*.

Sun, Y.: (2013), 'Fixed-smoothing Asymptotics in a Two-step GMM Framework'. Working Paper, Department of Economics, UCSD.

Sun, Y., P. C. B. Phillips, and S. Jin: (2008), 'Optimal Bandwidth Selection in Heteroskedasticity-Autocorrelation Robust Testing'. *Econometrica* **76**, 175–194.

Wooldridge, J. M. and H. White: (1988), 'Some invariance principles and central limit theorems for dependent heterogeneous processes'. *Econometric Theory* pp. 210–230.

# CHAPTER 2
# A Test of Parameter Instability Allowing for Change in the Moments of Explanatory Variables

## 2.1 Introduction

Most of the structural change literature addresses the instability of the parameters in the conditional model. The goal of this chapter is to develop a valid HAC-robust test when there is a shift in the mean and second moment of the explanatory variables. The break date, $\lambda_x T$, for the mean and second moment is assumed to be different from the break date $\lambda T$ for the regression parameters. The bootstrap approach in Hansen (2000) considers a setup where general form of the structural change in the marginal distribution of the explanatory variables are allowed. But Hansen (2000) assumes that the error term is a martingale difference sequence with respect to a certain information set. Thus serial correlation of the product of $x$ variables and the error, $u_t$, is not allowed. This can be a limitation particularly in the static time series regression model where the series $\{x_t u_t\}_{t=1,\ldots,T}$ often displays serial correlation. In this Chapter a valid HAC-robust approach for testing structural change in the regression slope/intercept parameters is developed. To ease the distribution theory a modified version of the standard set of high level conditions is introduced. To make the test robust to serial correlation and heteroskedasticity, a HAC estimator is used for constructing the test statistic and the fixed-$b$ theory developed in Kiefer and Vogelsang (2005) is applied to derive the asymptotic distribution. The limiting distributions of the statistics are pivotal. The rest of the chapter is organized as follows.

Section 2.2 lays out the basic set up of the problem. Section 2.3 derives the limiting distributions of the appropriate test statistics under the fixed-$b$ approach. Section 2.4 presents fixed-$b$ critical values for certain break point value and bandwidths. The finite sample properties of the test is examined in Monte Carlo simulation experiments. Section 2.5 summarizes and concludes. Proofs are collected in the Appendix.

## 2.2 Model of Structural Change and Preliminary Results

Suppose the univariate series $\{x_t\}$ has a mean shift at $t = \lambda_x T$. Denote

$$E(x_t) = \mu_{i(t)},$$

with

$$\mu_{i(t)} = \begin{cases} \mu_1 & \text{for } t \leq \lambda_x T \\ \mu_2 & \text{for } t \geq \lambda_x T + 1 \end{cases} . \tag{2.1}$$

The subscript $i(t)$ indicates a regime for time $t$ and $\mu_1$ is the mean in the first regime and $\mu_2$ is the mean in the second regime. Suppose $\lambda_x$, $\mu_1$, and $\mu_2$ are known.

Now consider a simple time series regression model with a structural break given by

$$\begin{aligned} y_t &= \alpha_1 D_t + \alpha_2(1 - D_t) + \beta_1 x_t D_t + \beta_2 x_t (1 - D_t) + u_t, \qquad (2.2) \\ D_t &= \mathbf{1}(t \leq \lambda T), \end{aligned}$$

where $x_t$ is *a* regressor, $\lambda \in (0, 1)$ is a hypothetical break point for the regression parameters, and $\mathbf{1}(\cdot)$ is the indicator function. For expositional simplicity let us suppose $\lambda T$ and $\lambda_x T$ are integer-valued. The regression model (2.2) implies that both regression parameters are subject to potential structural change (full structural change model). Consider the null hypothesis of no structural change in the slope parameter:

$$H_0 : \beta_1 = \beta_2 \tag{2.3}$$

The OLS estimators of $\alpha_1$, $\alpha_2$, $\beta_1$, and $\beta_2$ are given by

$$\begin{aligned}
\widehat{\beta}_1 &= \left( \sum_{t=1}^{\lambda T} (x_t - \overline{x}_1)^2 \right)^{-1} \left( \sum_{t=1}^{\lambda T} (x_t - \overline{x}_1)\, y_t \right), \\
\widehat{\beta}_2 &= \left( \sum_{t=\lambda T+1}^{T} (x_t - \overline{x}_2)^2 \right)^{-1} \left( \sum_{t=\lambda T+1}^{T} (x_t - \overline{x}_2)\, y_t \right), \\
\widehat{\alpha}_1 &= \overline{y}_1 - \widehat{\beta}_1 \overline{x}_1, \\
\widehat{\alpha}_2 &= \overline{y}_2 - \widehat{\beta}_2 \overline{x}_2,
\end{aligned} \tag{2.4}$$

where

$$\overline{y}_1 = \frac{1}{\lambda T} \sum_{t=1}^{\lambda T} y_t,\ \overline{y}_2 = \frac{1}{(1-\lambda)T} \sum_{t=\lambda T+1}^{T} y_t, \text{ and}$$
$$\overline{x}_1 = \frac{1}{\lambda T} \sum_{t=1}^{\lambda T} x_t,\ \overline{x}_2 = \frac{1}{(1-\lambda)T} \sum_{t=\lambda T+1}^{T} x_t.$$

Consider the case where $\lambda > \lambda_x$. The asymptotic distribution of the slope estimator can be obtained under a certain set of high level conditions. A typical set of conditions is:

**Assumption 1.** $T^{-1} \sum_{t=1}^{[rT]} x_t^2 \xrightarrow{p} r q_x^2$, uniformly in $r \in [0,1]$, and $q_x^2$ is strictly positive.

**Assumption 2.** $T^{-1/2} \sum_{t=1}^{[rT]} \begin{pmatrix} u_t \\ x_t u_t \end{pmatrix} \Rightarrow \widetilde{\Lambda} W(r)$, $r \in [0,1]$, where $\widetilde{\Lambda}\widetilde{\Lambda}' = \Sigma$, and $W(r)$ is a $2 \times 1$ standard Wiener process.

Assumptions 1 and 2 are the standard high level conditions used widely in the econometrics literature. Under these assumptions an immediate result is presented in the next Proposition without proof.

**Proposition 4.** *Suppose $\lambda$ is known and $\lambda > \lambda_x$. Under Assumptions 1 and 2 the OLS estimators* *in (2.4) have the following asymptotic distributions:*

$$T^{1/2}\left(\widehat{\beta}_1-\beta_1\right)\Rightarrow\left[\lambda q_x^2+\lambda_x\left(1-\frac{\lambda_x}{\lambda}\right)(\mu_1-\mu_2)^2\right]^{-1}\times\left(-\mu_2-\frac{\lambda_x}{\lambda}(\mu_1-\mu_2),\ 1\right)\cdot\widetilde{\Lambda}W(\lambda),$$

$$T^{1/2}\left(\widehat{\beta}_2-\beta_2\right)\Rightarrow\left[(1-\lambda)q_x^2\right]^{-1}\times(-\mu_2,\ 1)\cdot\widetilde{\Lambda}\left(W(1)-W(\lambda)\right).$$

For simplicity assume $q_x^2$ is known. To test the null hypothesis (2.3), one can consider using a robust Wald statistic

$$Wald\left(T_b\right)=T\left(\widehat{\beta}_1-\widehat{\beta}_2\right)\left(\widehat{avar}\left(T^{1/2}\left(\widehat{\beta}_1-\widehat{\beta}_2\right)\right)\right)^{-1}\left(\widehat{\beta}_1-\widehat{\beta}_2\right),$$

where $T_b=\lambda T$ denotes a hypothetical break date and

$$\widehat{avar}\left(T^{1/2}\left(\widehat{\beta}_1-\widehat{\beta}_2\right)\right)=$$

$$\lambda\left[\lambda q_x^2+\lambda_x\left(1-\frac{\lambda_x}{\lambda}\right)(\mu_1-\mu_2)^2\right]^{-2}\times\left(-\mu_2-\frac{\lambda_x}{\lambda}(\mu_1-\mu_2),\ 1\right)$$
$$\times\widehat{\Sigma}\times\left(-\mu_2-\frac{\lambda_x}{\lambda}(\mu_1-\mu_2),\ 1\right)'$$
$$+(1-\lambda)\left[(1-\lambda)q_x^2\right]^{-2}\times(-\mu_2,\ 1)\times\widehat{\Sigma}\times(-\mu_2,\ 1)',$$

where $\widehat{\Sigma}$ is a nonparametric kernel estimator of $\Sigma=\widetilde{\Lambda}\widetilde{\Lambda}'$ using a kernel $K(\cdot)$ and the bandwidth $M$:

$$\widehat{\Sigma}=T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}K\left(\frac{|t-s|}{M}\right)\widehat{v}_t\widehat{v}_s',$$

with $\widehat{v}_t=\begin{pmatrix}\widehat{u}_t\\ x_t\widehat{u}_t\end{pmatrix}$. Unfortunately, $Wald\left(T_b\right)$ does not have a pivotal asymptotic null distribution under fixed-$b$ asymptotics as long as $\mu_1$ is not equal to $\mu_2$ and neither is the supremum statistic. The reason is the vector, $\lambda\left[\lambda q_x^2+\lambda_x\left(1-\frac{\lambda_x}{\lambda}\right)(\mu_1-\mu_2)^2\right]^{-1}\times\left(-\mu_2-\frac{\lambda_x}{\lambda}(\mu_1-\mu_2),\ 1\right)$, is generally different from $(1-\lambda)\left[(1-\lambda)q_x^2\right]^{-1}\times(-\mu_2,\ 1)$

unless $\mu_1$ is equal to $\mu_2$.

In order to obtain a test statistic which has a pivotal fixed-$b$ limiting distribution, a modified regression equation and more general version of the high level conditions are introduced. The following two high-level conditions replace Assumptions 1 and 2.

**Assumption 1′.** $T^{-1}\sum_{t=1}^{[rT]}\left(x_t-\mu_{i(t)}\right)^2 \xrightarrow{p} \begin{cases} rs_1^2 \text{ uniformly in } r\in[0,\lambda_x] \\ \lambda_x s_1^2+(r-\lambda_x)\,s_2^2 \text{ uniformly in } r\in[\lambda_x,1] \end{cases}$
with $s_i^2>0$ for $i=1,2$.

Obviously under Assumption 1′

$$T^{-1}\sum_{t=1}^{[rT]}\left(\frac{x_t-\mu_{i(t)}}{s_{i(t)}}\right)^2 \xrightarrow{p} r \text{ uniformly in } r\in[0,1].$$

**Assumption 2′.** $T^{-1/2}\sum_{t=1}^{[rT]}\begin{pmatrix} \frac{1}{s_{i(t)}}u_t \\ \left(\frac{x_t-\mu_{i(t)}}{s_{i(t)}}\right)u_t \end{pmatrix} \Rightarrow \Lambda W(r), r\in[0,1]$, where $\Lambda\Lambda'=\Omega$, $W(r)$ is a $2\times 1$ standard Wiener process and $s_{i(t)}=s_1\mathbf{1}(t\leq\lambda_x T)+s_2\mathbf{1}(t\geq\lambda_x T+1)$.

When $\mu_1=\mu_2=\mu$ and $s_1^2=s_2^2=s^2$, Assumption 1′ is equivalent to Assumption 1 and Assumption 2′ is equivalent to Assumption 2. Under $\mu_1=\mu_2=\mu$ and $s_1^2=s_2^2=s^2$, Assumption 1′ implies $T^{-1}\sum_{t=1}^{[rT]}(x_t-\mu_t)^2 = T^{-1}\sum_{t=1}^{[rT]}(x_t-\mu)^2 \xrightarrow{p} rs^2$ and it is easy to show this is equivalent to Assumption 1 by defining $q_x^2=s^2+\mu^2$. Also, one can show the equivalence of Assumption 2′ and 2 through the relationship

$$\begin{pmatrix} s & 0 \\ \mu s & s \end{pmatrix}\Lambda=\widetilde{\Lambda}.$$

Note however that Assumption 1′ is more general than Assumption 1 since it allows the mean or probability limit of (centered and noncentered) sample second moment of $x_t$ to have a break. As a special case, if $x_t$ is a stationary process within each regime, then $s_i^2$ is the variance of $x_t$ process and Assumption 1′ is a robust version of Assumption 1 allowing for a shift of the variance. Assumption 2′ is also a more general version than Assumption 2. To see this why, suppose one has evidence of shift in $q_x^2$ (or $s_{i(t)}^2$). Then it would be hard to justify Assumption 2 as being still valid. For example, take the case where $x_t$ is stationary and conditional homoskedasticity holds, $E\left(u_t^2|x_t\right) = \sigma_u^2$. For a stationary variable, $E\left(x_t^2\right) = q_x^2$. Then the variance of $x_t u_t$, denoted by $\Gamma_0$, is $E\left(x_t^2 u_t^2\right) = E\left[E\left(x_t^2 u_t^2|x_t\right)\right] = \sigma_u^2 E\left(x_t^2\right) = \sigma_u^2 q_x^2$. Hence $\Gamma_0$ should have a structural break in general as long as there is a shift in $q_x^2$. This in turn implies that the long-run variance matrix $\Omega \equiv \Lambda\Lambda'$ also has structural break because $\Gamma_0$ is a component of the long run variance matrix. Therefore, it would not be appropriate to maintain Assumption 2 while allowing a shift in $q_x^2$ or $s_{i(t)}^2$ as in Assumption 1′. One needs to match Assumption 1′ with Assumption 2′ due to this reason. However, note that Assumption 2′ does not allow the long run variance to change in an arbitrary fashion. Assumption 2′ can only reflect the impact of change in $q_x^2$ or $s_{i(t)}^2$ on $\Omega$ in a particular way.

In next Section tests of parameter instability of $\alpha$ or(and) $\beta$ will be presented under a particular specification of breaks in $\left(\mu_{i(t)}, s_{i(t)}\right)$ and breaks in $(\alpha, \beta)$, which is as follows:

$\left(\mu_{i(t)}, s_{i(t)}\right)$ has a structural break at $\lambda_x T$ and $(\alpha, \beta)$ is allowed to have a break at $\lambda T$.

There may be an empirical application where a different specification is more relevant. For example, only $\mu_{i(t)}$ has structural break at $\lambda_x T$ and $(\alpha, \beta)$ are allowed to have break at $\lambda T$. But analysis for these other cases is not pursued in this chapter.

## 2.3 Asymptotic Results

Suppose $\left(\mu_{i(t)}, s_{i(t)}\right)$ has a break at $\lambda_x T$ and $(\alpha, \beta)$ is allowed to change at an unknown break date, $t = \lambda T$. Assume $\lambda_x$ and $\left(\mu_{i(t)}, s_{i(t)}\right)$ are given (known). Also, assume $\lambda \in [\epsilon, (\lambda_x - \epsilon)] \cup [(\lambda_x + \epsilon), 1 - \epsilon]$ following Andrews (1993). The admissible values for $\lambda$ is obtained by trimming the values at both ends of the sample period and the values in the neighborhood of $\lambda_x T$. It is necessarily implied that the break date for the regression parameters $(\alpha, \beta)$ should not be same as the break date for the moments. Recall the regression equation (2.2). Define $w_t = \frac{x_t}{s_{i(t)}}$. Without parameter instability, the equation can be written as

$$y_t = \alpha + \beta x_t + u_t = \alpha + \beta s_{i(t)} \left( \frac{x_t}{s_{i(t)}} \right) + u_t = \alpha + \beta \left( s_1 \mathbf{1} \left( t \leq \lambda_x T \right) + s_2 \mathbf{1} \left( t > \lambda_x T \right) \right) w_t + u_t$$

$$= \alpha + \beta s_1 \mathbf{1} \left( t \leq \lambda_x T \right) w_t + \beta s_2 \mathbf{1} \left( t > \lambda_x T \right) w_t + u_t.$$

Once we allow $(\alpha, \beta)$ to have a break at $\lambda T$, the above regression equation can be rewritten as

$$y_t = \alpha_1 \mathbf{1} \left( t \leq \lambda T \right) + \alpha_2 \mathbf{1} \left( t > \lambda T \right) + \left( \beta_1 \mathbf{1} \left( t \leq \lambda T \right) + \beta_2 \mathbf{1} \left( t > \lambda T \right) \right) s_1 \mathbf{1} \left( t \leq \lambda_x T \right) w_t$$

$$+ \left( \beta_1 \mathbf{1} \left( t \leq \lambda T \right) + \beta_2 \mathbf{1} \left( t > \lambda T \right) \right) s_2 \mathbf{1} \left( t > \lambda_x T \right) w_t + u_t.$$

Note that there are four interaction terms made by the two time indicator functions and one of these four interactions is identical to zero. For example, if $\lambda < \lambda_x$, then $\mathbf{1} \left( t \leq \lambda T \right) \times \mathbf{1} \left( t > \lambda_x T \right) = 0$. Depending on the relative magnitude of $\lambda$ and $\lambda_x$, one can rewrite the regression equation by reparametrization. For $\lambda < \lambda_x$

$$\begin{aligned} y_t = {} & \alpha_1 \mathbf{1} \left( t \leq \lambda T \right) + \alpha_2 \mathbf{1} \left( \lambda T < t \leq \lambda_x T \right) + \alpha_3 \mathbf{1} \left( t > \lambda_x T \right) \\ & + \gamma_1 w_t \mathbf{1} \left( t \leq \lambda T \right) + \gamma_2 w_t \mathbf{1} \left( \lambda T < t \leq \lambda_x T \right) + \gamma_3 w_t \mathbf{1} \left( t > \lambda_x T \right) + \varepsilon_t, \end{aligned} \tag{2.5}$$

where $\gamma_1 = \beta_1 s_1$, $\gamma_2 = \beta_2 s_1$, $\gamma_3 = \beta_2 s_2$, and $\alpha_2 = \alpha_3$. For $\lambda > \lambda_x$

$$\begin{aligned} y_t = {} & \alpha_1 \mathbf{1} \left( t \leq \lambda_x T \right) + \alpha_2 \mathbf{1} \left( \lambda_x T < t \leq \lambda T \right) + \alpha_3 \mathbf{1} \left( t > \lambda T \right) \\ & + \gamma_1 w_t \mathbf{1} \left( t \leq \lambda_x T \right) + \gamma_2 w_t \mathbf{1} \left( \lambda_x T < t \leq \lambda T \right) + \gamma_3 w_t \mathbf{1} \left( t > \lambda T \right) + \varepsilon_t, \end{aligned} \tag{2.6}$$

where $\gamma_1 = \beta_1 s_1$, $\gamma_2 = \beta_1 s_2$, $\gamma_3 = \beta_2 s_2$, and $\alpha_1 = \alpha_2$.

Denote as $W$ the $T \times 6$ matrix which collects the six explanatory variables: for (2.5), $\mathbf{1}(t \leq \lambda T), w_t\mathbf{1}(t \leq \lambda T), \mathbf{1}(\lambda T < t \leq \lambda_x T), w_t\mathbf{1}(\lambda T < t \leq \lambda_x T), \mathbf{1}(t > \lambda_x T), w_t\mathbf{1}(t > \lambda_x T)$ in this order and for (2.6), $\mathbf{1}(t \leq \lambda_x T), w_t\mathbf{1}(t \leq \lambda_x T), \mathbf{1}(\lambda_x T < t \leq \lambda T), w_t\mathbf{1}(\lambda_x T < t \leq \lambda T)$, $\mathbf{1}(t > \lambda T)$, and $w_t\mathbf{1}(t > \lambda T)$ in this order.

Notice that the extra parameter for $\alpha$ is introduced for each equation so that the regression model takes the form of full structural change model with three regimes.

The stability of $\alpha$ can be rewritten as $\alpha_1 = \alpha_2$ in (2.5) and $\alpha_2 = \alpha_3$ in (2.6). Likewise the stability of slope parameter $\beta$ can be rewritten as $\beta_1 = \beta_2$ in (2.5) and $\beta_2 = \beta_3$ in (2.6). The parameters in (2.5) and (2.6) are estimated by OLS and robust Wald statistics will be constructed based on these estimators. The OLS estimators in (2.5) are given by

$$\begin{aligned}
\widehat{\gamma}_1 &= \left(\sum_{t=1}^{\lambda T} (w_t - \overline{w}_1)^2\right)^{-1} \left(\sum_{t=1}^{\lambda T} (w_t - \overline{w}_1)\, y_t\right), \\
\widehat{\gamma}_2 &= \left(\sum_{t=\lambda T+1}^{\lambda_x T} (w_t - \overline{w}_2)^2\right)^{-1} \left(\sum_{t=\lambda T+1}^{\lambda_x T} (w_t - \overline{w}_2)\, y_t\right), \\
\widehat{\gamma}_3 &= \left(\sum_{t=\lambda_x T+1}^{T} (w_t - \overline{w}_3)^2\right)^{-1} \left(\sum_{t=\lambda_x T+1}^{T} (w_t - \overline{w}_3)\, y_t\right),
\end{aligned}$$

and

$$\begin{aligned}
\widehat{\alpha}_1 &= \overline{y}_1 - \widehat{\gamma}_1 \overline{w}_1, \\
\widehat{\alpha}_2 &= \overline{y}_2 - \widehat{\gamma}_2 \overline{w}_2, \\
\widehat{\alpha}_3 &= \overline{y}_3 - \widehat{\gamma}_3 \overline{w}_3,
\end{aligned}$$

where

$$\begin{array}{rlcrl}
\overline{w}_1 &= \frac{1}{\lambda T} \sum_{t=1}^{\lambda T} \frac{x_t}{s_1} & & \overline{y}_1 &= \frac{1}{\lambda T} \sum_{t=1}^{\lambda T} y_t \\
\overline{w}_2 &= \frac{1}{\lambda_x T - \lambda T} \sum_{t=\lambda T+1}^{\lambda_x T} \frac{x_t}{s_1} & \text{and} & \overline{y}_2 &= \frac{1}{\lambda_x T - \lambda T} \sum_{t=\lambda T+1}^{\lambda_x T} y_t \\
\overline{w}_3 &= \frac{1}{T - \lambda_x T} \sum_{t=\lambda_x T+1}^{T} \frac{x_t}{s_2} & & \overline{y}_3 &= \frac{1}{T - \lambda_x T} \sum_{t=\lambda_x T+1}^{T} y_t
\end{array}.$$

Defining $\overline{x}_1, \overline{x}_2$, and $\overline{x}_3$ similarly as above immediately gives $\overline{w}_1 = \frac{\overline{x}_1}{s_1}$, $\overline{w}_2 = \frac{\overline{x}_2}{s_1}$, and $\overline{w}_3 =$

$\frac{\overline{x}_3}{s_2}$. One can easily write down the expression for the OLS estimators $(\widetilde{\gamma}_1, \widetilde{\gamma}_2, \widetilde{\gamma}_3, \widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{\alpha}_3)$ and all the other quantities when $\lambda \in [(1+\epsilon)\lambda_x, 1-\epsilon]$. The next Proposition presents the asymptotic limits of the OLS estimators.

**Proposition 5.** *Suppose* $\lambda \in [\epsilon, (1-\epsilon)\lambda_x]$. *Denote* $\Lambda = \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix}$. *Under Assumptions 1′ and 2′, as* $T \to \infty$ *the OLS estimators in (2.5) have the following limits:*

$$\begin{aligned}
T^{1/2}(\widehat{\gamma}_1 - \gamma_1) &\Rightarrow \Lambda_2 \frac{W(\lambda)}{\lambda}, \\
T^{1/2}(\widehat{\alpha}_1 - \alpha_1) &\Rightarrow \left(s_1\Lambda_1 - \frac{\mu_1}{s_1}\Lambda_2\right)\frac{W(\lambda)}{\lambda}, \\
T^{1/2}(\widehat{\gamma}_2 - \gamma_2) &\Rightarrow \Lambda_2 \frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda}, \text{ and} \\
T^{1/2}(\widehat{\alpha}_2 - \alpha_2) &\Rightarrow \left(s_1\Lambda_1 - \frac{\mu_1}{s_1}\Lambda_2\right)\frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda}.
\end{aligned}$$

*Suppose* $\lambda \in [(1+\epsilon)\lambda_x, 1-\epsilon]$. *Then the OLS estimators in (2.6) have following limits:*

$$\begin{aligned}
T^{1/2}(\widetilde{\gamma}_2 - \gamma_2) &\Rightarrow \frac{1}{\lambda - \lambda_x}\Lambda_2 (W(\lambda) - W(\lambda_x)), \\
T^{1/2}(\widetilde{\alpha}_2 - \alpha_2) &\Rightarrow \left(s_2\Lambda_1 - \frac{\mu_2}{s_2}\Lambda_2\right)\frac{W(\lambda) - W(\lambda_x)}{\lambda - \lambda_x}, \\
T^{1/2}(\widetilde{\gamma}_3 - \gamma_3) &\Rightarrow \frac{1}{1-\lambda}\Lambda_2 (W(1) - W(\lambda)), \text{ and} \\
T^{1/2}(\widetilde{\alpha}_3 - \alpha_3) &\Rightarrow \left(s_2\Lambda_1 - \frac{\mu_2}{s_2}\Lambda_2\right)\frac{W(1) - W(\lambda)}{1-\lambda}.
\end{aligned}$$

**Proof:** See the Appendix.

### 2.3.1 Stability of $\beta$

Consider the null hypothesis

$$H_0 : \beta \text{ is stable.} \tag{2.7}$$

This null hypothesis implies $\gamma_1 = \gamma_2$ in (2.5) and $\gamma_2 = \gamma_3$ in (2.6). From Proposition 10, for $\lambda \in [\epsilon, \lambda_x - \epsilon]$,

$$T^{1/2}\left(\widehat{\gamma}_2 - \widehat{\gamma}_1\right) \Rightarrow \Lambda_2 \left( \frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda} - \frac{W(\lambda)}{\lambda} \right) \tag{2.8}$$
$$\sim N\left(0, \frac{\lambda_x}{\lambda(\lambda_x - \lambda)} \Lambda_2 \Lambda_2'\right),$$

and for $\lambda \in [\lambda_x + \epsilon, 1 - \epsilon]$,

$$T^{1/2}\left(\widetilde{\gamma}_3 - \widetilde{\gamma}_2\right) \Rightarrow \Lambda_2 \left( \frac{W(1) - W(\lambda)}{1 - \lambda} - \frac{W(\lambda) - W(\lambda_x)}{\lambda - \lambda_x} \right) \tag{2.9}$$
$$\sim N\left(0, \frac{1 - \lambda_x}{(1-\lambda)(\lambda - \lambda_x)} \Lambda_2 \Lambda_2'\right).$$

**Test Statistic $T_1$**

Denote $T_b = \lambda T$ and define robust Wald statistics:

$$Wald_1^{\beta}(T_b) = \frac{T\left(\widehat{\gamma}_2 - \widehat{\gamma}_1\right)^2}{\widehat{\Lambda_2 \Lambda_2'}} \text{ for } T_b \in [\epsilon T, (\lambda_x - \epsilon) T] \equiv \Xi_1 \text{ and} \tag{2.10}$$

$$Wald_2^{\beta}(T_b) = \frac{T\left(\widetilde{\gamma}_3 - \widetilde{\gamma}_2\right)^2}{\widehat{\Lambda_2 \Lambda_2'}} \text{ for } T_b \in [(\lambda_x + \epsilon) T, (1 - \epsilon) T] \equiv \Xi_2, \tag{2.11}$$

where $\widehat{\Lambda_2 \Lambda_2'}$ is a nonparametric kernel HAC estimator given by

$$\widehat{\Lambda_2 \Lambda_2'} = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right) \widehat{v}_t \widehat{v}_s', \tag{2.12}$$
$$\text{with } \widehat{v}_t = \left(\frac{x_t - \mu_{i(t)}}{s_{i(t)}}\right) \widehat{u}_t.$$

The HAC estimator $\widehat{\Lambda_2 \Lambda_2'}$ in $Wald_1^{\beta}(T_b)$ is computed using the residuals $\widehat{u}_t$ from the regression of equation (2.5) and $\widehat{\Lambda_2 \Lambda_2'}$ in $Wald_2^{\beta}(T_b)$ is computed with the residuals $\widehat{u}_t$ from the regression of equation (2.6). Under the assumption of a fixed bandwidth ratio this

HAC estimator can be rewritten as a function of partial sum processes (see Kiefer and Vogelsang (2005)),

$$\widehat{S}_{[rT]}^{\beta} = \sum_{t=1}^{[rT]} \widehat{v}_t = \sum_{t=1}^{[rT]} \left( \frac{x_t - \mu_{i(t)}}{s_{i(t)}} \right) \widehat{u}_t.$$

The next Proposition presents the limit of this (scaled) partial sum process.

**Proposition 6.** *Under Assumptions 1′ and 2′, as $T \to \infty$, the limit of the partial sum process is given by*

*When $0 < \lambda < \lambda_x$,*

$$T^{1/2}\widehat{S}_{[rT]}^{\beta} \Rightarrow \begin{cases} \Lambda_2 \left( W(r) - \frac{r}{\lambda} W(\lambda) \right) \text{ for } 0 \le r \le \lambda, \\ \Lambda_2 \left( W(r) - W(\lambda) - \frac{r-\lambda}{\lambda_x - \lambda} \left( W(\lambda_x) - W(\lambda) \right) \right) \text{ for } \lambda \le r \le \lambda_x, \\ \Lambda_2 \left( W(r) - W(\lambda_x) - \frac{r-\lambda_x}{1-\lambda_x} \left( W(1) - W(\lambda_x) \right) \right) \text{ for } \lambda_x \le r \le 1, \end{cases}$$

*and*

*When $\lambda_x < \lambda < 1$,*

$$T^{1/2}\widehat{S}_{[rT]}^{\beta} \Rightarrow \begin{cases} \Lambda_2 \left( W(r) - \frac{r}{\lambda_x} W(\lambda_x) \right) \text{ for } 0 \le r \le \lambda_x, \\ \Lambda_2 \left( W(r) - W(\lambda_x) - \frac{r-\lambda_x}{\lambda - \lambda_x} \left( W(\lambda) - W(\lambda_x) \right) \right) \text{ for } \lambda_x \le r \le \lambda, \\ \Lambda_2 \left( W(r) - W(\lambda) - \frac{r-\lambda}{1-\lambda} \left( W(1) - W(\lambda) \right) \right) \text{ for } \lambda \le r \le 1. \end{cases}$$

**Proof:** See the Appendix.

Define

$$\begin{aligned} H_1 &\equiv H_1 \left( r, \lambda, \lambda_x \right) \\ &= \left( W_1(r) - \frac{r}{\lambda} W_1(\lambda) \right) \cdot \mathbf{1}(r \le \lambda) \\ &+ \left( W_1(r) - W_1(\lambda) - \frac{r-\lambda}{\lambda_x - \lambda} \left( W_1(\lambda_x) - W_1(\lambda) \right) \right) \cdot \mathbf{1}(\lambda < r \le \lambda_x) \\ &+ \left( W_1(r) - W_1(\lambda_x) - \frac{r-\lambda_x}{1-\lambda_x} \left( W_1(1) - W_1(\lambda_x) \right) \right) \cdot \mathbf{1}(\lambda_x < r \le 1), \end{aligned}$$

and

$$H_2 \equiv H_2 \left( r, \lambda, \lambda_x \right)$$

$$= \left( W_1(r) - \frac{r}{\lambda_x} W_1(\lambda_x) \right) \cdot \mathbf{1}(r \leq \lambda_x)$$

$$+ \left( W_1(r) - W_1(\lambda_x) - \frac{r - \lambda_x}{\lambda - \lambda_x} \left( W_1(\lambda) - W_1(\lambda_x) \right) \right) \cdot \mathbf{1}(\lambda_x < r \leq \lambda)$$

$$+ \left( W_1(r) - W_1(\lambda) - \frac{r - \lambda}{1 - \lambda} \left( W_1(1) - W_1(\lambda) \right) \right) \cdot \mathbf{1}(\lambda < r \leq 1),$$

where $W_1(\cdot)$ is a 1-dimensional Wiener process.

**Theorem 4.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1′ and 2′, as $T \to \infty$, the limits under the null hypothesis in (2.7) is given by*

$$Wald_1^{\beta}\left(T_b\right) \overset{H_0}{\Rightarrow} \frac{\left( \frac{W_1(\lambda_x) - W_1(\lambda)}{\lambda_x - \lambda} - \frac{W_1(\lambda)}{\lambda} \right)^2}{\mathbf{P}\left(b, H_1\right)} \equiv Wald_1^{\infty}\left(\lambda, \lambda_x\right),$$

*and,*

$$Wald_2^{\beta}\left(T_b\right) \overset{H_0}{\Rightarrow} \frac{\left( \frac{W_1(1) - W_1(\lambda)}{1 - \lambda} - \frac{W_1(\lambda) - W_1(\lambda_x)}{\lambda - \lambda_x} \right)^2}{\mathbf{P}\left(b, H_2\right)} \equiv Wald_2^{\infty}\left(\lambda, \lambda_x\right),$$

*The definition of $\mathbf{P}\left(b, H_1\right)$ and $\mathbf{P}\left(b, H_2\right)$ can be found in Cho and Vogelsang (2014).*

**Proof:** See the Appendix.

Finally, define a test statistic $T_1$:

$$T_1 = \max\left( \max_{T_b \in \Xi_1} Wald_1^{\beta}\left(T_b\right), \max_{T_b \in \Xi_2} Wald_2^{\beta}\left(T_b\right) \right). \tag{2.13}$$

This statistic can be used when the break date is unknown. Its limit is given by

$$\max\left( \sup_{\lambda \in [\epsilon, (1-\epsilon)\lambda_x]} Wald_1^{\infty}\left(\lambda, \lambda_x\right), \sup_{\lambda \in [(1+\epsilon)\lambda_x, 1-\epsilon]} Wald_2^{\infty}\left(\lambda, \lambda_x\right) \right).$$

**Test Statistic $T_1^{(F)}$**

Alternatively, the inference in this setup can be based on different HAC estimators defined in Cho and Vogelsang (2014). These alternative HAC estimators are given by

$$\Upsilon_1 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right) \widetilde{v}_t^{(1)} \widetilde{v}_s^{(1)\prime}, \tag{2.14}$$

$$\widetilde{v}_t^{(1)} = \begin{pmatrix} \left(\frac{x_t-\mu_1}{s_1}\right) \mathbf{1}\left(t \leq \lambda T\right) \\ \left(\frac{x_t-\mu_1}{s_1}\right) \mathbf{1}\left(\lambda T < t \leq \lambda_x T\right) \\ \left(\frac{x_t-\mu_2}{s_2}\right) \mathbf{1}\left(\lambda_x T \leq t \leq T\right) \end{pmatrix} \widehat{u}_t,$$

where $\lambda < \lambda_x$, and

$$\Upsilon_2 = T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right) \widetilde{v}_t^{(2)} \widetilde{v}_s^{(2)\prime}, \tag{2.15}$$

$$\widetilde{v}_t^{(2)} = \begin{pmatrix} \left(\frac{x_t-\mu_1}{s_1}\right) \mathbf{1}\left(t \leq \lambda_x T\right) \\ \left(\frac{x_t-\mu_2}{s_2}\right) \mathbf{1}\left(\lambda_x T < t \leq \lambda T\right) \\ \left(\frac{x_t-\mu_2}{s_2}\right) \mathbf{1}\left(\lambda T \leq t \leq T\right) \end{pmatrix} \widehat{u}_t,$$

where $\lambda > \lambda_x$. Define robust Wald statistics based on the above HAC estimators:

$$Wald_1^{(F),\beta}\left(T_b\right) = \frac{T\left(\widehat{\gamma}_2 - \widehat{\gamma}_1\right)^2}{D_1 Q_1^{-1} \Upsilon_1 Q_1^{-1} D_1'} \text{ for } T_b \in \left[\epsilon T, \left(\lambda_x - \epsilon\right) T\right] \equiv \Xi_1, \tag{2.16}$$

$$Wald_2^{(F),\beta}\left(T_b\right) = \frac{T\left(\widetilde{\gamma}_3 - \widetilde{\gamma}_2\right)^2}{D_2 Q_2^{-1} \Upsilon_2 Q_2^{-1} D_2'} \text{ for } T_b \in \left[\left(\lambda_x + \epsilon\right) T, \left(1 - \epsilon\right) T\right] \equiv \Xi_2, \tag{2.17}$$

where

$$D_1 \equiv (1, \ -1) \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$Q_1 = \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{\lambda T}(w_t - \overline{w}_1)^2 & 0 & 0 \\ 0 & \frac{1}{T}\sum_{t=\lambda T+1}^{\lambda_x T}(w_t - \overline{w}_2)^2 & 0 \\ 0 & 0 & \frac{1}{T}\sum_{t=\lambda_x T+1}^{T}(w_t - \overline{w}_3)^2 \end{pmatrix},$$

$$D_2 \equiv (1, \ -1) \times \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and}$$

$$Q_2 = \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{\lambda_x T}(w_t - \overline{w}_1)^2 & 0 & 0 \\ 0 & \frac{1}{T}\sum_{t=\lambda_x T+1}^{\lambda T}(w_t - \overline{w}_2)^2 & 0 \\ 0 & 0 & \frac{1}{T}\sum_{t=\lambda T+1}^{T}(w_t - \overline{w}_3)^2 \end{pmatrix}.$$

**Theorem 5.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1′ and 2′, as $T \to \infty$, the limits under the null hypothesis in (2.7) is given by*

$$Wald_1^{(F),\beta}(T_b) \overset{H_0}{\Rightarrow} \frac{\left(\frac{W_1(\lambda_x)-W_1(\lambda)}{\lambda_x-\lambda} - \frac{W_1(\lambda)}{\lambda}\right)^2}{\mathbf{P}(b, H_3)} \equiv Wald_1^{(F),\infty}(\lambda, \lambda_x),$$

*and,*

$$Wald_2^{(F),\beta}(T_b) \overset{H_0}{\Rightarrow} \frac{\left(\frac{W_1(1)-W_1(\lambda)}{1-\lambda} - \frac{W_1(\lambda)-W_1(\lambda_x)}{\lambda-\lambda_x}\right)^2}{\mathbf{P}(b, H_4)} \equiv Wald_2^{(F),\infty}(\lambda, \lambda_x),$$

*where the processes $H_3$ and $H_4$ are defined as*

$$H_3 \equiv H_3(r, \lambda, \lambda_x)$$

$$= \frac{1}{\lambda}\left(W_1(r) - \frac{r}{\lambda}W_1(\lambda)\right) \cdot \mathbf{1}(r \le \lambda)$$

$$-\frac{1}{\lambda_x - \lambda}\left(W_1(r) - W_1(\lambda) - \frac{r-\lambda}{\lambda_x - \lambda}\left(W_1(\lambda_x) - W_1(\lambda)\right)\right) \cdot \mathbf{1}(\lambda < r \le \lambda_x),$$

*and*

$$H_4 \equiv H_4\left(r,\lambda,\lambda_x\right)$$

$$= \frac{1}{\lambda-\lambda_x}\left(W_1(r)-W_1(\lambda_x)-\frac{r-\lambda_x}{\lambda-\lambda_x}\left(W_1(\lambda)-W_1(\lambda_x)\right)\right)\cdot\mathbf{1}(\lambda_x<r\leq\lambda)$$
$$-\frac{1}{1-\lambda_x}\left(W_1(r)-W_1(\lambda)-\frac{r-\lambda}{1-\lambda}\left(W_1(1)-W_1(\lambda)\right)\right)\cdot\mathbf{1}(\lambda<r\leq 1).$$

The test statistic for testing the null hypothesis in (2.7) is simply given by

$$\begin{aligned}
T_1^{(F)} &= \max\left(\max_{T_b\in\Xi_1} Wald_1^{(F),\beta}\left(T_b\right), \max_{T_b\in\Xi_2} Wald_2^{(F),\beta}\left(T_b\right)\right)\\
&\Rightarrow \max\left(\sup_{\lambda\in[\epsilon,(1-\epsilon)\lambda_x]} Wald_1^{(F),\infty}\left(\lambda,\lambda_x\right), \sup_{\lambda\in[(1+\epsilon)\lambda_x,1-\epsilon]} Wald_2^{(F),\infty}\left(\lambda,\lambda_x\right)\right)\\
&\equiv T^{(F),\infty}\left(\epsilon,\lambda_x\right).
\end{aligned}$$

### 2.3.2 Stability of $\alpha$

Consider the null hypothesis

$$H_0 : \alpha \text{ is stable.} \tag{2.18}$$

With a hypothetical break point $\lambda T$, the break in $\alpha$ may occur at $\lambda T$. This implies $\alpha_1 = \alpha_2$ in (2.5) and $\alpha_2 = \alpha_3$ in (2.6) under the null hypothesis (2.18). From Proposition 10, for $\lambda \in [\epsilon, \lambda_x - \epsilon]$,

$$\begin{aligned}
T^{1/2}\left(\widehat{\alpha}_2-\widehat{\alpha}_1\right) &\Rightarrow \left(s_1\Lambda_1-\frac{\mu_1}{s_1}\Lambda_2\right)\left(\frac{W\left(\lambda_x\right)-W\left(\lambda\right)}{\lambda_x-\lambda}-\frac{W\left(\lambda\right)}{\lambda}\right) \\
&\sim N\left(0,\frac{\lambda_x}{\lambda\left(\lambda_x-\lambda\right)}\left(s_1,\ -\frac{\mu_1}{s_1}\right)\Lambda\Lambda'\begin{pmatrix} s_1 \\ -\frac{\mu_1}{s_1}\end{pmatrix}\right),
\end{aligned} \tag{2.19}$$

and for $\lambda \in [\lambda_x + \epsilon, 1 - \epsilon]$,

$$T^{1/2}\left(\widetilde{\alpha}_3 - \widetilde{\alpha}_2\right) \Rightarrow \left(s_2\Lambda_1 - \frac{\mu_2}{s_2}\Lambda_2\right)\left(\frac{W(1) - W(\lambda)}{1-\lambda} - \frac{W(\lambda) - W(\lambda_x)}{\lambda - \lambda_x}\right) \tag{2.20}$$
$$\sim N\left(0, \frac{1-\lambda_x}{(1-\lambda)(\lambda - \lambda_x)}\left(s_2, \ -\frac{\mu_2}{s_2}\right)\Lambda\Lambda'\begin{pmatrix} s_2 \\ -\frac{\mu_2}{s_2}\end{pmatrix}\right).$$

**Test Statistic $T_2$**

Denote $T_b = \lambda T$ and define robust Wald statistics:

$$Wald_1^{\alpha}(T_b) = \frac{T\left(\widehat{\alpha}_2 - \widehat{\alpha}_1\right)^2}{\left(s_1, \ -\frac{\mu_1}{s_1}\right) \times \widehat{\Lambda\Lambda'} \times \left(s_1, \ -\frac{\mu_1}{s_1}\right)'} \text{ for } T_b \in [\epsilon T, (\lambda_x - \epsilon)T] \equiv \Xi_1 \text{ and} \tag{2.21}$$

$$Wald_2^{\alpha}(T_b) = \frac{T\left(\widetilde{\alpha}_3 - \widetilde{\alpha}_2\right)^2}{\left(s_2, \ -\frac{\mu_2}{s_2}\right) \times \widehat{\Lambda\Lambda'} \times \left(s_2, \ -\frac{\mu_2}{s_2}\right)'} \text{ for } T_b \in [(\lambda_x + \epsilon)T, (1-\epsilon)T] \equiv \Xi_2, \tag{2.22}$$

where $\widehat{\Lambda\Lambda'}$ is a nonparametric kernel HAC estimator given by

$$\widehat{\Lambda\Lambda'} = T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T} K\left(\frac{|t-s|}{M}\right)\widehat{\xi}_t\widehat{\xi}_s', \tag{2.23}$$
$$\text{with } \widehat{\xi}_t = \begin{pmatrix} \frac{1}{s_{i(t)}} \\ \frac{x_t - \mu_{i(t)}}{s_{i(t)}} \end{pmatrix}\widehat{u}_t.$$

As before, the HAC estimator $\widehat{\Lambda\Lambda'}$ in $Wald_1^{\alpha}(T_b)$ is computed using the residuals $\widehat{u}_t$ from the regression of equation (2.5) and $\widehat{\Lambda\Lambda'}$ in $Wald_2^{\alpha}(T_b)$ is computed with the residuals $\widehat{u}_t$ from the regression of equation (2.6). Under the assumption of fixed bandwidth ratio, this HAC estimator can be rewritten as a function of partial sum processes (see Kiefer and Vogelsang (2005)),

$$\widehat{S}_{[rT]}^{\alpha} = \sum_{t=1}^{[rT]}\widehat{\xi}_t = \sum_{t=1}^{[rT]}\begin{pmatrix} \frac{1}{s_{i(t)}} \\ \frac{x_t - \mu_{i(t)}}{s_{i(t)}} \end{pmatrix}\widehat{u}_t.$$

The next Proposition presents the limit of the (scaled) partial sum processes.

**Proposition 7.** *Under Assumptions 1′ and 2′, as $T \to \infty$, the limit of the partial sum process is given by*

$$\text{When } 0 < \lambda < \lambda_x,$$

$$T^{1/2}\widehat{S}^{\alpha}_{[rT]} \Rightarrow \begin{cases} \Lambda\left(W(r) - \frac{r}{\lambda}W(\lambda)\right) \text{ for } 0 \le r \le \lambda, \\ \Lambda\left(W(r) - W(\lambda) - \frac{r-\lambda}{\lambda_x-\lambda}\left(W(\lambda_x) - W(\lambda)\right)\right) \text{ for } \lambda \le r \le \lambda_x, \\ \Lambda\left(W(r) - W(\lambda_x) - \frac{r-\lambda_x}{1-\lambda_x}\left(W(1) - W(\lambda_x)\right)\right) \text{ for } \lambda_x \le r \le 1, \end{cases}$$

*and*

$$\text{When } \lambda_x < \lambda < 1,$$

$$T^{1/2}\widehat{S}^{\alpha}_{[rT]} \Rightarrow \begin{cases} \Lambda\left(W(r) - \frac{r}{\lambda_x}W(\lambda_x)\right) \text{ for } 0 \le r \le \lambda_x, \\ \Lambda\left(W(r) - W(\lambda_x) - \frac{r-\lambda_x}{\lambda-\lambda_x}\left(W(\lambda) - W(\lambda_x)\right)\right) \text{ for } \lambda_x \le r \le \lambda, \\ \Lambda\left(W(r) - W(\lambda) - \frac{r-\lambda}{1-\lambda}\left(W(1) - W(\lambda)\right)\right) \text{ for } \lambda \le r \le 1. \end{cases}$$

**Proof:** See the Appendix.

The next Theorem presents the limit of the statistics. The limit is the same as in Theorem 8.

**Theorem 6.** *Let $\lambda \in (0,1)$ and $b \in (0,1]$ be given. Suppose $M = bT$. Then under Assumptions 1′ and 2′, as $T \to \infty$, the limits under the null hypothesis in (2.7) is given by*

$$Wald_1^{\alpha}\left(T_b\right) \overset{H_0}{\Rightarrow} Wald_1^{\infty}\left(\lambda, \lambda_x\right),$$

*and,*

$$Wald_2^{\alpha}\left(T_b\right) \overset{H_0}{\Rightarrow} Wald_2^{\infty}\left(\lambda, \lambda_x\right).$$

*The definition of $\mathbf{P}\left(b, H_1\right)$ and $\mathbf{P}\left(b, H_2\right)$ can be found in Cho and Vogelsang (2014).*

**Proof:** See the Appendix.

Using Theorem 9, the test statistic $T_2$ defined below has the following limit:

$$\begin{aligned} T_2 &= \max\left(\max_{T_b \in \Xi_1} Wald_1^{\alpha}\left(T_b\right), \max_{T_b \in \Xi_2} Wald_2^{\alpha}\left(T_b\right)\right) \\ &\Rightarrow \max\left(\sup_{\lambda \in [\epsilon, (1-\epsilon)\lambda_x]} Wald_1^{\infty}\left(\lambda, \lambda_x\right), \sup_{\lambda \in [(1+\epsilon)\lambda_x, 1-\epsilon]} Wald_2^{\infty}\left(\lambda, \lambda_x\right)\right). \end{aligned}$$

**Test Statistic** $T_2^{(F)}$

Alternative statistic can be derived by a similar way as before. Construct HAC estimators $\Upsilon_3$, $\Upsilon_4$ using $\widetilde{\xi}_t^{(1)}$ and $\widetilde{\xi}_t^{(2)}$ respectively where

$$\widetilde{\xi}_t^{(1)} = \begin{pmatrix} \frac{1}{s_1}\mathbf{1}\left(t \leq \lambda T\right) \\ \left(\frac{x_t - \mu_1}{s_1}\right)\mathbf{1}\left(t \leq \lambda T\right) \\ \frac{1}{s_1}\mathbf{1}\left(\lambda T < t \leq \lambda_x T\right) \\ \left(\frac{x_t - \mu_1}{s_1}\right)\mathbf{1}\left(\lambda T < t \leq \lambda_x T\right) \\ \frac{1}{s_2}\mathbf{1}\left(\lambda_x T \leq t \leq T\right) \\ \left(\frac{x_t - \mu_2}{s_2}\right)\mathbf{1}\left(\lambda_x T \leq t \leq T\right) \end{pmatrix} \times \widehat{u}_t,$$

and

$$\widetilde{\xi}_t^{(2)} = \begin{pmatrix} \frac{1}{s_1}\mathbf{1}\left(t \leq \lambda_x T\right) \\ \left(\frac{x_t - \mu_1}{s_1}\right)\mathbf{1}\left(t \leq \lambda_x T\right) \\ \frac{1}{s_2}\mathbf{1}\left(\lambda_x T < t \leq \lambda T\right) \\ \left(\frac{x_t - \mu_2}{s_2}\right)\mathbf{1}\left(\lambda_x T < t \leq \lambda T\right) \\ \frac{1}{s_2}\mathbf{1}\left(\lambda T \leq t \leq T\right) \\ \left(\frac{x_t - \mu_2}{s_2}\right)\mathbf{1}\left(\lambda T \leq t \leq T\right) \end{pmatrix} \times \widehat{u}_t.$$

Define Wald statistics

$$Wald_1^{(F),\alpha}\left(T_b\right) = \frac{T\left(\widehat{\alpha}_2 - \widehat{\alpha}_1\right)^2}{\left(s_1, \ -\frac{\mu_1}{s_1}\right) \times D_3\left(\frac{W'W}{T}\right)^{-1}\Upsilon_3\left(\frac{W'W}{T}\right)^{-1}D_3' \times \left(s_1, \ -\frac{\mu_1}{s_1}\right)'} \tag{2.24}$$

for $T_b \in [\epsilon T, (\lambda_x - \epsilon) T] \equiv \Xi_1$, and

$$Wald_2^{(F),\alpha}(T_b) = \frac{T(\widetilde{\alpha}_3 - \widetilde{\alpha}_2)^2}{\left(s_2, \ -\frac{\mu_2}{s_2}\right) \times D_4 \left(\frac{W'W}{T}\right)^{-1} \Upsilon_4 \left(\frac{W'W}{T}\right)^{-1} D_4' \times \left(s_2, \ -\frac{\mu_2}{s_2}\right)'} \tag{2.25}$$

for $T_b \in [(\lambda_x + \epsilon) T, (1-\epsilon) T] \equiv \Xi_2$, where

$$D_3 \equiv (I_2, \ -I_2) \times (I_4, \ \mathbf{0}_{4\times 1}),$$
$$D_4 \equiv (I_2, \ -I_2) \times (\mathbf{0}_{4\times 1}, \ I_4).$$

Finally,

$$\begin{aligned} T_2^{(F)} &\equiv \max\left(\max_{T_b\in\Xi_1} Wald_1^{(F),\alpha}(T_b), \max_{T_b\in\Xi_2} Wald_2^{(F),\alpha}(T_b)\right) \\ &\Rightarrow \max\left(\sup_{\lambda\in[\epsilon,(1-\epsilon)\lambda_x]} Wald_1^{(F),\infty}(\lambda,\lambda_x), \sup_{\lambda\in[(1+\epsilon)\lambda_x,1-\epsilon]} Wald_2^{(F),\infty}(\lambda,\lambda_x)\right). \end{aligned}$$

The asymptotic limits of $Wald_1^{(F),\alpha}(T_b)$, $Wald_2^{(F),\alpha}(T_b)$ and $T_2^{(F)}$ are the same as $Wald_1^{(F),\infty}(\lambda,\lambda_x)$, $Wald_1^{(F),\infty}(\lambda,\lambda_x)$ and $T^{(F),\infty}(\epsilon,\lambda_x)$ respectively in previous Theorem.

## 2.4 Simulations

In this Section the finite sample properties of the test are examined via Monte Carlo simulation. The data generating process (DGP) is given by

$$y_t = \alpha_1 \mathbf{1}(t \le \lambda T) + \alpha_2 \mathbf{1}(t > \lambda T) + \beta_1 \mathbf{1}(t \le \lambda T) x_t + \beta_2 \mathbf{1}(t > \lambda T) x_t + \varepsilon_t,$$
$$x_t = s_{i(t)} u_t + \mu_{i(t)}, \ \ u_t = \rho u_{t-1} + \eta_t,$$
$$\varepsilon_t = 0.5\varepsilon_{t-1} + \nu_t,$$

where $\eta_t \sim N(0, \sigma_\eta^2)$ and $v_t \sim N(0,1)$ are independent. The values of $\alpha_1$ and $\alpha_2$ are set to zero. The mean and variance of $x_t$ are allowed to have a single break at $\lambda_x = 0.4$. Three specifications on the mean and variance of $x_t$ are considered.

$$\begin{aligned}
&\text{specification 0: } \left(\mu_1,\ \mu_2, s_1^2,\ s_2^2\right) = (0,\ 0,\ 1,\ 1) \\
&\text{specification 1: } \left(\mu_1,\ \mu_2,\ s_1^2,\ s_2^2\right) = (0,\ 0.5,\ 1,\ 1.5) \\
&\text{specification 2: } \left(\mu_1,\ \mu_2,\ s_1^2,\ s_2^2\right) = (0,\ 1,\ 1,\ 2) \\
&\text{specification 3: } \left(\mu_1,\ \mu_2,\ s_1^2,\ s_2^2\right) = (0,\ 2,\ 1,\ 4)
\end{aligned}$$

Specification 0 implies there is no break in the mean and variance. When the true values of $(\lambda_x,\ \mu_1,\ \mu_2,\ s_1^2,\ s_2^2)$ are known, one can recover $u_t$ from $x_t$ and use $u_t$ to construct a HAC estimator. So the break in mean and variance has no effect on inference and the empirical rejection probabilities should be same across the specifications. The values of $\left(\rho,\ \sigma_\eta^2\right)$ are chosen so that the variance of $x_t$ becomes 1 for the first regime (before $\lambda_x T$). The selected sets of those values are (0.5, 0.75) and (0.8, 0.36). The DGP with $\left(\rho,\ \sigma_\eta^2\right) = (0.8,\ 0.36)$ has more persistent autocovariance function than the other and has the larger long run variance of $u_t$.

To examine the size property, the case where $\beta_1 = \beta_2 = 1$ is considered. To learn about power, the following values for $(\lambda^0,\ \beta_1,\ \beta_2)$ are considered: $(0.2,\ 1,\ 1.5)$, $(0.25,\ 1,\ 1.5)$, where $\lambda^0$ denotes the true break point for the regression parameters. The value of the trimming parameter to be used in this experiment is 0.1. Based on this particular value of trimming parameter and the value of $\lambda_x$, the admissible set of $\lambda$ is $[0.1,\ 0.4 - 0.1] \sqcup [0.4 + 0.1,\ 1 - 0.1]$. To check against the case where the true value does not belong to this admissible set, extra values of $\lambda$ is considered: $\lambda = 0.4$. With $\epsilon = 0.1$ and $\lambda_x = 0.4$ and the Bartlett kernel being used, the 95% fixed-$b$ critical values for the statistic $T_1$ are 188.64 for $b = 0.1$ and 717.15 for $b = 0.5$. For $T_1^{(F)}$ with $\epsilon = 0.1$ and $\lambda_x = 0.4$ the 95% fixed-$b$ critical values are 35.11 for $b = 0.1$ and 162.07 for $b = 0.5$. Table 2.1 reports the results for the simulation experiments for $T = 100$ and 300 and for three different tests: $SupW^{(F)}$ test in Cho and Vogelsang (2014), and $T_1$ and $T_1^{(F)}$ tests proposed in this chapter.

Table 2.1 shows that the test results for $T_1$ and $T_1^{(F)}$ are invariant to the break in the moments of the explanatory variable. This is, as mentined earlier, because the true value of $\left(\lambda_x, \mu_1, \mu_2, s_1^2, s_2^2\right)$ is assumed to be observable and these two statistics are numerically the same across different values of $\left(\lambda_x, \mu_1, \mu_2, s_1^2, s_2^2\right)$. There are several points conveyed by this table. First, in the absence of break in the moments, $SupW^{(F)}$ test displays better power property and comparable size property over the other two tests. But when there exists a break in the moments, the size and power of $SupW^{(F)}$ test would be sensitive to the magnitude of the change in the moments and the size property does not get better as $T$ increases. When there is change in the moments and $T_1^{(F)}$ test is used for the inference, the size distortion is relatively small and the rejection frequency is not so much sensitive to the persistence of the underlying process. However, this test has poor power compared to $SupW^{(F)}$ test. This is because for given a value of $b$, $T_1^{(F)}$ uses bigger effective bandwidth for estimating the variance matrix in each regime. The test based on $T_1$ is seen to be dominated by $T_1^{(F)}$ test in terms of the size and power when $T$ is 300.

## 2.5 Summary and Conclusions

This chapter proposes an inference procedure for testing stability of regression parameters allowing for a single break in the mean and second moments of the $x$ variable. The mean and second moments are assumed to have a break at $\lambda_x T$. The break point for the regression parameter ($\lambda T$) should be different from $\lambda_x T$. The analysis focuses on a simple linear regression model but the proposed test can be generalized to a multiple linear regression model. A new set of high level conditions are introduced which incorporates the possibility of change in the mean and second moments of the $x$ variable. Under fixed-$b$ asymptotics, the limiting distribution of the robust Wald statistic is pivotal so the critical values can be simulated and used for conducting inference. The simulation results in this chapter show there is substantial size distortion in finite samples. This is not surprising because the three regimes induced by two different types of break points make even smaller sample size for each regime. Also, whether the main results will be still valid when $\lambda_x$ and the moments are unknown and need to be estimated should be a straightforward direction of future research.

Table 2.1: Size and Power in Finite Samples, T=100, 300, $\lambda_x = 0.4$, $\epsilon = .1$, Bartlett kernel

| | No break in $(\mu, s^2)$ | | | | $\mu_1 = 0, \mu_2 = 0.5$ $s_1^2 = 1, s_2^2 = 1.5$ | | | | $\mu_1 = 0, \mu_2 = 1$ $s_1^2 = 1, s_2^2 = 2$ | | | | $\mu_1 = 0, \mu_2 = 2$ $s_1^2 = 1, s_2^2 = 4$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\left(\rho,\ \sigma_\eta^2\right)$ | (.5, .75) | | (.8, .36) | | (.5, .75) | | (.8, .36) | | (.5, .75) | | (.8, .36) | | (.5, .75) | | (.8, .36) | |
| $b$ | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 |
| | | | | | | | | $SupW^{(F)}$ Test | | | | | | | | |
| T=100 | | | | | | | | | | | | | | | | |
| Size | 13.3 | 11.9 | 23.0 | 20.7 | 16.5 | 14.1 | 25.3 | 21.1 | 24.6 | 19.4 | 30.8 | 26.7 | 63.8 | 50.8 | 58.8 | 49.6 |
| Power | | | | | | | | | | | | | | | | |
| $\lambda^0 = .2$ | 27.9 | 22.1 | 33.1 | 28.2 | 46.8 | 36.7 | 45.3 | 37.2 | 67.5 | 52.8 | 60.5 | 50.7 | 98.1 | 86.7 | 93.6 | 83.6 |
| .25 | 28.1 | 22.6 | 33.4 | 28.7 | 48.6 | 38.0 | 46.5 | 38.5 | 69.0 | 55.0 | 61.8 | 51.4 | 98.3 | 88.6 | 94.1 | 84.5 |
| .4 | 26.3 | 20.5 | 32.0 | 26.7 | 47.2 | 36.9 | 46.2 | 38.2 | 70.4 | 56.5 | 62.4 | 52.8 | 98.4 | 93.2 | 94.3 | 86.8 |
| T=300 | | | | | | | | | | | | | | | | |
| Size | 8.4 | 7.9 | 14.4 | 11.9 | 14.8 | 13.3 | 18.6 | 16.0 | 29.9 | 25.1 | 30.9 | 25.6 | 93.0 | 83.6 | 84.8 | 75.0 |
| Power | | | | | | | | | | | | | | | | |
| $\lambda^0 = .2$ | 41.5 | 34.8 | 38.8 | 32.9 | 76.0 | 59.8 | 67.6 | 55.5 | 93.4 | 75.1 | 86.8 | 71.9 | 100 | 96.0 | 99.9 | 95.6 |
| .25 | 41.8 | 34.1 | 38.7 | 32.5 | 78.3 | 63.6 | 69.4 | 57.4 | 95.2 | 81.0 | 88.6 | 74.7 | 100 | 97.3 | 99.8 | 96.6 |
| .4 | 37.8 | 30.6 | 36.3 | 29.3 | 80.0 | 68.0 | 70.6 | 58.3 | 97.1 | 88.8 | 90.6 | 79.8 | 100 | 99.8 | 99.9 | 99 |

| | $T_1$ Test | | | | $T_1^{(F)}$ Test | | | |
|---|---|---|---|---|---|---|---|---|
| $\left(\rho,\ \sigma_\eta^2\right)$ | (.5, .75) | | (.8, .36) | | (.5, .75) | | (.8, .36) | |
| $b$ | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 | 0.1 | 0.5 |
| T=100 | | | | | | | | |
| Size | 20.1 | 18.7 | 45.7 | 41.2 | 12.0 | 10.7 | 13.7 | 12.3 |
| Power | | | | | | | | |
| $\lambda^0 = .2$ | 25.6 | 21.6 | 49.6 | 43.2 | 20.3 | 18.9 | 17.3 | 15.8 |
| .25 | 25.2 | 21.9 | 49.3 | 43.6 | 19.5 | 19.0 | 17.2 | 16.4 |
| .4 | 20.1 | 18.7 | 45.7 | 41.2 | 12.0 | 10.7 | 13.7 | 12.3 |
| T=300 | | | | | | | | |
| Size | 9.5 | 9.4 | 20.9 | 18.9 | 8.6 | 7.8 | 11.8 | 11.1 |
| Power | | | | | | | | |
| $\lambda^0 = .2$ | 25.1 | 20.5 | 31.9 | 25.8 | 25.0 | 25.3 | 22.2 | 22.1 |
| .25 | 25.3 | 19.8 | 31.4 | 25.0 | 24.6 | 24.9 | 22.0 | 21.6 |
| .4 | 9.5 | 9.4 | 20.9 | 18.9 | 8.6 | 7.8 | 11.8 | 11.1 |

APPENDIX

# Appendix for Chapter 2

**Proof of Proposition 10:** The proof is only provided for $\widehat{\gamma}_2$. One can easily prove the rest of the results.

$$\begin{aligned}\widehat{\gamma}_2 &= \left(\sum_{t=\lambda T+1}^{\lambda_x T} (w_t - \overline{w}_2)^2\right)^{-1} \left(\sum_{t=\lambda T+1}^{\lambda_x T} (w_t - \overline{w}_2)\, y_t\right) \\ &= \gamma_2 + \left(\sum_{t=\lambda T+1}^{\lambda_x T} \left(\frac{x_t - \overline{x}_2}{s_1}\right)^2\right)^{-1} \left(\sum_{t=\lambda T+1}^{\lambda_x T} \left(\frac{x_t - \overline{x}_2}{s_1}\right) u_t\right).\end{aligned}$$

Since $p\lim \overline{x}_2 = p\lim \frac{1}{(\lambda_x - \lambda)T} \sum_{t=\lambda T+1}^{\lambda_x T} x_t = \mu_1$, it follows under Assumptions 1′ and 2′,

$$T^{1/2}\left(\widehat{\gamma}_2 - \gamma_2\right) \Rightarrow \Lambda_2 \left(\frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda}\right).$$

**Proof of Proposition 11:** Proof is only provided for $r \in [\lambda, \lambda_x]$ when $0 < \lambda < \lambda_x$.

$$\widehat{S}^{\beta}_{[rT]} = \sum_{t=1}^{[rT]} \left(\frac{x_t - \mu_1}{s_1}\right) \widehat{u}_t = \sum_{t=\lambda T+1}^{[rT]} \left(\frac{x_t - \mu_1}{s_1}\right) \left(y_t - \widehat{\alpha}_2 - \widehat{\gamma}_2 \frac{x_t}{s_1}\right)$$

$$= (\alpha_2 - \widehat{\alpha}_2) \sum_{t=\lambda T+1}^{[rT]} \left(\frac{x_t - \mu_1}{s_1}\right) + (\gamma_2 - \widehat{\gamma}_2) \sum_{t=\lambda T+1}^{[rT]} \left(\frac{x_t - \mu_1}{s_1}\right) \frac{x_t}{s_1} + \sum_{t=\lambda T+1}^{[rT]} \left(\frac{x_t - \mu_1}{s_1}\right) u_t.$$

So,

$$\begin{aligned}T^{-1/2}\widehat{S}^{\beta}_{[rT]} &\Rightarrow -\Lambda_2 \left(\frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda}\right) \times (r - \lambda) + \Lambda_2 \left(W(r) - W(\lambda)\right) \\ &= \Lambda_2 \left(W(r) - W(\lambda) - \frac{r - \lambda}{\lambda_x - \lambda} \left(W(\lambda_x) - W(\lambda)\right)\right).\end{aligned}$$

**Proof of Theorem 8:** The results immediately follows from equation (2.8), (2.9), Proposition 3 and the transformation

$$\Lambda_2 W(\cdot) \overset{d}{=} A W_1(\cdot),$$

where $A$ is the positive constant satisfying $\Lambda_2\Lambda_2' = A^2$.

**Proof of Proposition 12:** Proof is only provided for $r \in [\lambda, \lambda_x]$ when $0 < \lambda < \lambda_x$.

$$\widehat{S}^{\alpha}_{[rT]} = \sum_{t=1}^{[rT]} \widehat{\xi}_t = \sum_{t=1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} \widehat{u}_t$$

$$= \sum_{t=\lambda T+1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} \widehat{u}_t = \sum_{t=\lambda T+1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} \left( y_t - \widehat{\alpha}_2 - \widehat{\gamma}_2 \frac{x_t}{s_1} \right)$$

$$= (\alpha_2 - \widehat{\alpha}_2) \sum_{t=\lambda T+1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} + (\gamma_2 - \widehat{\gamma}_2) \sum_{t=\lambda T+1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} \frac{x_t}{s_1} + \sum_{t=\lambda T+1}^{[rT]} \begin{pmatrix} \frac{1}{s_1} \\ \frac{x_t - \mu_1}{s_1} \end{pmatrix} u_t.$$

Hence

$$T^{-1/2} \widehat{S}^{\alpha}_{[rT]} \Rightarrow -(r-\lambda) \begin{pmatrix} \frac{1}{s_1} \\ 0 \end{pmatrix} \left( s_1, \ -\frac{\mu_1}{s_1} \right) \Lambda \left( \frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda} \right)$$

$$-(r-\lambda) \begin{pmatrix} \frac{\mu_1}{s_1^2} \\ 1 \end{pmatrix} \Lambda_2 \left( \frac{W(\lambda_x) - W(\lambda)}{\lambda_x - \lambda} \right) + \Lambda \left( W(r) - W(\lambda) \right)$$

$$= \Lambda \left( W(r) - W(\lambda) - \frac{r-\lambda}{\lambda_x - \lambda} \left( W(\lambda_x) - W(\lambda) \right) \right).$$

**Proof of Theorem 9:** The results immediately follows from equation (2.19), (2.20), Proposition 4 and the transformations

$$\left( s_1, \ -\frac{\mu_1}{s_1} \right) \Lambda W(\cdot) \stackrel{d}{=} B_1 W_1(\cdot),$$
$$\left( s_2, \ -\frac{\mu_2}{s_2} \right) \Lambda W(\cdot) \stackrel{d}{=} B_2 W_1(\cdot),$$

where the constants $B_1$ and $B_2$ satisfies

$$\left(s_1, \ -\frac{\mu_1}{s_1}\right) \Lambda\Lambda' \left(s_1, \ -\frac{\mu_1}{s_1}\right)' = B_1^2,$$
$$\left(s_2, \ -\frac{\mu_2}{s_2}\right) \Lambda\Lambda' \left(s_2, \ -\frac{\mu_2}{s_2}\right)' = B_2^2.$$

# REFERENCES

REFERENCES

Andrews, D. W. K.: (1993), 'Tests for Parameter Instability and Structural Change with Unknown Change Point'. *Econometrica* **61**, 821–856.

Cho, C. K. and T. J. Vogelsang: (2014), 'Fixed-*b* Inference for Testing Structural Change in a Time Series Regression'. Working paper, Department of Economics, Michigan State University.

Hansen, B. E.: (2000), 'Testing for structural change in conditional models'. *Journal of Econometrics* **97**(1), 93–115.

Kiefer, N. M. and T. J. Vogelsang: (2005), 'A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests'. *Econometric Theory* **21**, 1130–1164.

# CHAPTER 3
# A Test of the Null of Integer Integration against the Alternative of Fractional Integration

## 3.1 Introduction

Chapter 2 proposes a test of the null hypothesis of "integer integration" against the alternative of fractional integration. More precisely, the null is that the series is either $I(0)$ or $I(1)$, while the alternative is that it is $I(d)$ with $0 < d < 1$. The null of integer integration is rejected in favor of the alternative of fractional integration if the KPSS test rejects the null of $I(0)$ and a unit root test rejects the null of $I(1)$. A new unit root test is used as the second part of the testing procedure, which is a lower-tailed KPSS test based on first differences of the data, but other unit root tests like the ADF test could also have been used. This two-part testing procedure will be called the "Double-KPSS" test because it consists of two steps, but it should be pointed out that the test is treated as one test and is evaluated for its properties (consistency and finite sample size and power) as such.

The KPSS test of Kwiatkowski et al. (1992) was originally suggested as a test of the null of (short memory) stationarity against the alternative of a unit root. Conversely, standard unit root tests like the Dickey-Fuller tests, the augmented Dickey-Fuller (ADF) test of Said and Dickey (1984) or the Phillips-Perron test of Phillips and Perron (1988) were viewed as tests of the null of a unit root against the alternative of short-memory stationarity. So if the KPSS test rejected but the unit root test did not, the conclusion was that the series had a unit root. If the unit root test rejected but the KPSS test did not, the conclusion was that the series was short-memory stationary. If neither test rejected, the conclusion was that the data were not informative enough to decide whether the series was $I(0)$ or $I(1)$. However, if both tests rejected, there was in some sense a contradiction.

This apparent contradiction can be resolved by considering a wider class of processes, specifically long-memory processes. The leading example considered in this chapter, is the $I(d)$ process (with $0 < d < 1$) of Adenstedt (1974), Granger and Joyeux (1980), and Hosking (1981). Since both the KPSS test and unit root tests have power against long-memory alternatives, the "double rejection" outcome can be taken as evidence that the series has long memory, as opposed to being either $I(0)$ or $I(1)$. This is not a novel observation. However, the approach in this chapter is novel in its consideration of the double-testing procedure as a single test, and its investigation of the size and power properties of this test. In this regard, the basic observation is that if the nominal size of each of the two tests is set to 5%, the double test also has size of 5% asymptotically. For example, if the DGP is $I(0)$, then asymptotically the KPSS test will reject with probability 5% while the unit root test will reject with probability one, while if the DGP is $I(1)$ the converse will occur. So whether the data are $I(0)$ or $I(1)$, the probability of rejection of the double test is asymptotically 5%.

The practical issue to be faced is to what extent one can be reasonably sure that the double rejection outcome is due to fractional integration, as opposed to size distortions of the test under the $I(0)$ or $I(1)$ null. For example, Caner and Kilian (2001) and Müller (2005) have shown that the KPSS test has large size distortions if the DGP is $AR(1)$ with autoregressive coefficient near unity. Conversely, Dejong et al. (1992), Phillips and Perron (1988) and Vogelsang and Wagner (2013), among others, have found that the Dickey-Fuller test and its variants can have large size distortions, especially if the DGP is $ARIMA(0, 1, 1)$ with moving average root near (negative) unity. This does not imply that the Double KPSS test will suffer from large size distortions in either of these cases, since the cases for which the KPSS test has large size distortions correspond to cases in which the unit root test may have low power, and conversely. However, it does argue for a careful investigation of the size and power properties of the new test in finite samples.

As noted above, the specific unit root test used in this chapter is a lower-tail KPSS test based on the data in differences. The KPSS unit root test suggested by Shin and Schmidt (1992) and Breitung (2002) was considered. However, as shown by Lee and Amsler (1997), the KPSS unit root test is not consistent against $I(d)$ alternatives with $1/2 < d < 1$. One might also consider using the ADF test, but this test is known to have low power against $I(d)$ alternatives (e.g. Diebold and Rudebusch (1991), Hassler and Wolters (1994)), and there is also the practical consideration that it is easier to prove the consistency of our test against $I(d)$ alternatives for all $d$ between zero and one than it is for the ADF test. In simulation, it makes little difference whether the new test or the ADF test is used.

The consistency of the Double KPSS test depends on the consistency of the KPSS test and of the unit root test proposed in this chapter, and these in turn depend on the number of lags used in the estimation of the long-run variance going to infinity, but more slowly than sample size. Under this assumption a single critical value for each test (for each significance level) obtains, and these will be referred to as the "standard asymptotics" critical values. They do not depend on the kernel used to estimate the long-run variance or on the bandwidth (so long as the number of lags behaves as assumed above). However, following Kiefer and Vogelsang (2002a, 2002b, 2005), this chapter also considers "fixed-$b$ asymptotics," where $b$, defined as the ratio of the number of lags to the sample size, is constant as the sample size grows. The fixed-$b$ critical values depend on the kernel and on the value of $b$, and there is evidence in many settings that they yield tests with smaller size distortions than the critical values based on the standard asymptotics.

The main theoretical contribution is that the consistency of the Double-KPSS test is proved against $I(d)$ for all $d$ between zero and one. For the KPSS test, this can be shown using existing results except for the case of $d = 1/2$, so the divergence of the statistic for $d = 1/2$ is proven in this chapter. For the new unit root test, its asymptotic distribution is established for $d = 0$, $0 < d < 1/2$ and $1/2 < d < 1$, and it is proven that the statistic converges to zero in probability when $d = 1/2$. Besides these theoretical results, this chapter contains substantial simulation results to show the extent to which this testing procedure is likely to be useful in finite samples.

The plan of the chapter is as follows. Section 3.2 gives the definitions and basic properties of stationary short memory, long memory and unit root processes, and explicitly states the testing procedure. Section 3.3 gives the asymptotic results, using the standard asymptotics. The asymptotic limits of the two component tests are derived and consistency of the two-part test is proved. Section 3.4 presents the fixed-$b$ asymptotics. Section 3.5 presents the results of simulations which explore the finite sample properties of the new test. Section 3.6 summarizes and concludes. Finally, an Appendix gives some proofs and technical details.

## 3.2 Setup and Assumptions

The data is assumed to be generated by the DGP:

$$y_t = \mu + \epsilon_t, \; t = 1, 2, ..., T. \tag{2.1}$$

That is, non-zero level of the $y_t$ series is allowed, but not trend. Allowing for trend would not change the basic principles of the research in this chapter, but it would change the asymptotics.

### 3.2.1 Null Hypothesis

Under the null hypothesis $\{\epsilon_t\}_{t=1}^{\infty}$ is either a stationary short memory process or a unit root process. That is, either $\{\epsilon_t\}_{t=1}^{\infty}$ itself is a stationary short memory process or it is a cumulation of a short memory process.

Let $\{z_t\}_{t=1}^{\infty}$ be a time series with zero mean, and let $Z_t = \sum_{j=1}^{t} z_j$ be its partial sum.

$\{z_t\}_{t=1}^{\infty}$ is said to be a short-memory process if it satisfies the following two conditions.

**Assumption N1**

$$\sigma^2 = \lim_{T\to\infty} T^{-1} E\left(Z_T^2\right) \text{ exists and is nonzero.} \tag{2.2}$$

**Assumption N2**

$$\forall r \in [0,1],\ T^{-1/2} Z_{[rT]} \Rightarrow \sigma W(r), \tag{2.3}$$

where $[rT]$ denotes the integer part of $rT$, $\Rightarrow$ means weak convergence, and $W(r)$ is the standard Wiener process.

In addition to Assumptions N1 and N2, further regularity conditions are necessary for the consistency of HAC (heteroskedasticity and autocorrelation consistent) estimators. Examples of such conditions can be found in Andrews (1991), Newey and West (1987), De Jong and Davidson (2000), Jansson (2002), and Hansen (1992). It is implicitly assumed that one or more of these sets of conditions hold, so that the HAC estimators that appear in our test statistics are consistent.

Unit root processes are the other class of DGP which belongs to the null hypothesis. A time series is said to be a unit root process if its first difference is a short memory process. Equivalently, a cumulation of a short memory process is a unit root process. That is, $Z_t$ is a unit root process if

$$(1-L)\, Z_t \equiv z_t \sim \text{short memory process.} \tag{2.4}$$

### 3.2.2 Alternative Hypothesis

Under the alternative hypothesis, $\{\epsilon_t\}_{t=1}^{\infty}$ is a fractionally integrated process. Specifically, consider the alternative that $\epsilon_t$ follows an $I(d)$ process with $0 < d < 1$ :

$$(1-L)^d \epsilon_t = u_t,\ u_t \sim \text{i.i.d Normal}(0, \sigma_u^2), \tag{2.5}$$

The class of $I(d)$ processes with $0 < d < \frac{1}{2}$ has been widely used in econometrics to represent long memory processes[1]. More generally, a stationary process is said to have long memory if

$$\lim_{n\to\infty} \sum_{j=-n}^{n} \gamma_j = \infty, \tag{2.6}$$

where $\gamma_j$ is the autocovariance at lag $j$. Lo (1991) uses the following form of autocovariance function as a definition of a long-range dependent (long memory) process.

$$\gamma_j \sim \left\{ \begin{array}{cc} j^{2d-1}L(j) \text{ for } d \in (0, \frac{1}{2}) \text{ or} & \\ & \text{as } j \to \infty \\ -j^{2d-1}L(j) \text{ for } d \in (-\frac{1}{2}, 0) & \end{array} \right\}, \tag{2.7}$$

where $L(j)$ is a slowly varying function[2] at infinity. This form of autocovariance function includes the autocovariance function of the $I(d)$ process with $0 < d < \frac{1}{2}$, and is more general in the sense that it would accommodate the case that $u_t$ in (2.5) is a short memory but not necessarily an i.i.d. process. However, it does not accommodate the case of $\frac{1}{2} \leq d < 1$.

This chapter considers the $I(d)$ process with i.i.d. innovations as in equation (2.5), which was analyzed by Granger and Joyeux (1980) and Hosking (1981). When $\left(-\frac{1}{2} <\right) d < \frac{1}{2}$, $d \neq 0$, the process is a stationary long memory process, while it is a nonstationry long memory process for $\frac{1}{2} \leq d < 1$. For $d > -\frac{1}{2}$, the process is invertible and has infinite

[1] For more comprehensive treatment of this topic, see Giraitis et al. (2012).

[2] Lo (1991, page 1286): A function $L(x)$ is said to be slowly varying at infinity if $\lim_{t\to\infty} L(tx)/L(t) = 1$. An example is $\log x$.

order moving average representation:

$$\epsilon_t = (1-L)^{-d} u_t = \sum_{j=0}^{\infty} b_j u_{t-j},\ b_j = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \tag{2.8}$$

and when $d < \frac{1}{2}$ it has infinite order AR representation:

$$(1-L)^d \epsilon_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j} = u_t, a_0 = 1,\ a_j = \frac{\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} \text{ for } j \geq 1. \tag{2.9}$$

Also, for $-\frac{1}{2} < d < \frac{1}{2}$, $d \neq 0$ the process has a slowly decaying autocovariance function given by

$$\gamma_k = \frac{\sigma_u^2 \Gamma(1-2d)\Gamma(k+d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)} \sim ck^{2d-1} \text{ as } k \to \infty \text{ for some constant } c, \tag{2.10}$$

and this autocovariance function satisfies (2.6) and (2.7). Hosking (1981) provided further results with ARMA$(p,q)$ innovations.

To establish the asymptotic results in this chapter, an invariance principle under the alternative hypothesis is needed. Davydov (1970) and Sowell (1990) provide an invariance principle for the fractionally integrated processes with i.i.d. innovations. Lee and Schmidt (1996) use the result in Sowell (1990), replacing his $r$th moment-condition by a normality assumption. Lo (1991) bases his asymptotic analysis upon the result of Taqqu (1975), assuming stationarity and Gaussianity of the long-memory process. More recently, Qiu and Lin (2011) proved an invariance principle for the fractionally integrated process with strong near-epoch dependent innovations, which is the most general functional central limit theorem for fractionally integrated processes currently available. The analysis in this chapter will focus on the $I(d)$ process with normal i.i.d. innovations, following Lee and Schmidt (1996). Specifically, this chapter will use the functional central limit theorem appearing in Sowell (1990) which is restated in Lee and Schmidt (1996), as follows.[3]

Suppose $\{z_t\}_{t=1}^{\infty}$ is generated by (2.5) with $-\frac{1}{2} < d < \frac{1}{2}$ and let $\sigma_T^2 = var\left(Z_T\right)$. From Sowell (1990),

$$\sigma_T^2 = \sigma_u^2 \cdot \frac{\Gamma\left(1-2d\right)}{\left(1+2d\right)\Gamma\left(1+d\right)\Gamma\left(1-d\right)} \times \left[\frac{\Gamma\left(1+d+T\right)}{\Gamma\left(T-d\right)} - \frac{\Gamma\left(1+d\right)}{\Gamma\left(-d\right)}\right] \tag{2.11}$$

and as $T \to \infty$,

$$\frac{\sigma_T^2}{T^{1+2d}} \to \sigma_u^2 \cdot \frac{\Gamma\left(1-2d\right)}{\left(1+2d\right)\Gamma\left(1+d\right)\Gamma\left(1-d\right)} \equiv \omega_d^2. \tag{2.12}$$

Also, Sowell (1990) gives the following invariance principle for the fractionally integrated processes with $-\frac{1}{2} < d < \frac{1}{2}$ :

$$\sigma_T^{-1} Z_{[rT]} \Rightarrow W_d(r), \tag{2.13}$$

or equivalently

$$T^{-(d+1/2)} Z_{[rT]} \Rightarrow \omega_d W_d(r), \tag{2.14}$$

where $W_d(r)$ is the fractional Brownian motion of Mandelbrot and Van Ness (1968), which is defined as

$$W_d(r) = \frac{1}{\Gamma\left(1+d\right)} \cdot \int_0^1 \left(r-s\right)^d dW(s) \tag{2.15}$$

### 3.2.3 Test Statistics and the Rejection Rule

As above, the model is:

$$y_t = \mu + \epsilon_t, \ t = 1, 2, ..., T. \tag{2.16}$$

The null hypothesis is: $H_0$ : $\epsilon_t$ is a stationary short-memory process or a unit root process with short-memory innovations, and the alternative hypothesis is: $H_1$ : $\epsilon_t$ is an $I(d)$ process with $0 < d < 1$ and with normal i.i.d. innovations.

[3]This result could presumably be generalized under the more general conditions in Qiu and Lin (2011). The technical details involved would be orthogonal to the main points of this paper.

New rejection rule will be based on two statistics. The first statistic, denoted by $K_1$, should have a non-degenerate limit only when the innovation $\epsilon_t$ is stationary short-memory, and should diverge when the innovation has a unit root or is a long memory process. Conversely, the second statistic, denoted by $K_2$, should have a non-degenerate asymptotic distribution only when the innovation has a unit root, and should converge to zero when the $\epsilon_t$ is a stationary short memory or when it is a long memory ($0 < d < 1$) process. If one can find two such statistics, $K_1$ and $K_2$, the following rejection rule will asymptotically have size of 5% :

$$\text{Reject } H_0 \text{ if } K_1 > cv^1_{0.95} \text{ and } K_2 < cv^2_{0.05},$$

where $cv^1_{0.95}$ is the upper 5% critical value from the asymptotic distribution of $K_1$ when the error term is a short memory process, and $cv^2_{0.05}$ is the lower 5% critical value from the asymptotic distribution of $K_2$ when the error term is a unit root process. Now two specific statistics ($K_1$ and $K_2$) are proposed to make this procedure operational.

The first statistic ($K_1$) is the KPSS statistic (KPSS, 1992) which is defined as

$$\widehat{\eta}_\mu = \frac{T^{-2}\sum_{t=1}^{T} S_t^2}{s^2(l)}, \tag{2.17}$$

where $s^2(l)$ is a HAC estimator. This statistic is constructed using the OLS residuals $\{e_j\}_{j=1}^{T}$ from equation (2.16). More specifically,

$$\begin{aligned} e_t &= y_t - \overline{y} = y_t - \frac{1}{T}\sum_{j=1}^{T} y_j, \;\; S_t = \sum_{j=1}^{t} e_j, \\ s^2(l) &= \widehat{\gamma}_0 + 2\sum_{s=1}^{l} w(s,l)\widehat{\gamma}_s, \end{aligned} \tag{2.18}$$

where $w(s,l) = 1 - \frac{s}{l+1}$ (Bartlett kernel) , $\widehat{\gamma}_s = \frac{1}{T}\sum_{j=s+1}^{T} e_j e_{j-s}$. Note that with the Bartlett kernel the number of lags, $l$, determines the maximum lag of the sample autocovariances considered for estimating the long run variance of the error term. It is assumed that $l \to \infty$ but $l/T \to 0$ as $T \to \infty$. This assumption characterizes the "traditional asymptotics" and is made to ensure the consistency of the test. Later Section will make some comparisions to the "fixed-$b$" asymptotics that arise when $b \equiv \frac{l+1}{T}$ has a fixed non-zero limit.

Section 3.3 collects existing asymptotic results for $\widehat{\eta}_\mu$. The only case in which $\widehat{\eta}_\mu$ has a non-degenerate limiting distribution occurs when the innovation follows a short-memory process. Under all the other cases, i.e. a unit root process and the processes described by our alternative hypothesis, $\widehat{\eta}_\mu$ diverges to infinity.

Also, a statistic $K_2$ is needed that has a non-degenerate limit only under the unit root innovation processes. Shin and Schmidt (1992) and Breitung (2002) considered a slightly modified KPSS statistic, $\frac{l}{T}\widehat{\eta}_\mu$, which converges to a non-degenerate limiting distribution under the unit root innovation process and goes to zero under short-memory error processes. But as shown by Lee and Amsler (1997), this statistic cannot distinguish $I(1)$ processes from $I(d)$ processes with $\frac{1}{2} < d < 1$. This chapter therefore suggest an alternative statistic which can distinguish $I(1)$ processes from $I(d)$ processes with $\frac{1}{2} \leq d < 1$ (nonstationary long-memory processes) as well as from $I(d)$ with $0 < d < \frac{1}{2}$ (stationary long-memory processes) and stationary short-memory processes.

Consider the differenced model,

$$\triangle y_t = \triangle \epsilon_t = \triangle e_t. \tag{2.19}$$

The second statistic $K_2$ will be the KPSS statistic based on the differenced data $\{\triangle y_t\}_{t=2}^{T}$. The statistic is given by

$$\widehat{\eta}_\mu^d = \frac{T^{-2}\sum_{t=2}^{T} \widetilde{S}_t^2}{\widetilde{s}^2(l)}, \tag{2.20}$$

where

$$\widetilde{S}_t = \sum_{j=2}^{t} \triangle y_j = \sum_{j=2}^{t} \triangle \epsilon_j = \epsilon_t - \epsilon_1, \tag{2.21}$$

$$\widetilde{s}^2(l) = \widetilde{\gamma}_0 + 2\sum_{s=1}^{l} w(s,l)\widetilde{\gamma}_s,$$

with $\widetilde{\gamma}_s = \frac{1}{T-1}\sum_{t=s+2}^{T} \Delta y_t \Delta y_{t-s} = \frac{1}{T-1}\sum_{t=s+2}^{T} \Delta\epsilon_t \Delta\epsilon_{t-s}$.

It will be shown that this statistic has a non-degenerate distribution in the case of a unit root, while it converges to zero under stationary short memory or under $I(d)$ with $0 < d < 1$. Note that this will be a lower-tail test. Other unit root test could have considered, notably variants of the Dickey-Fuller test. The simulations will also compare the results with different unit root tests. From a technical point of view, $\widehat{\eta}_\mu^d$ is attractive because the proof of the consistency of the test is relatively straightforward.

## 3.3 Asymptotic Results

This Section discusses the asymptotic distributions of the $\widehat{\eta}_\mu$ and $\widehat{\eta}_\mu^d$ statistics when $\epsilon_t$ is $I(0)$, $I(d)$ with $0 < d < 1$, and $I(1)$. The asymptotic theory is established under the assumption that, as $T \to \infty$, $l \to \infty$ but $l/T \to 0$, where $l$ is the number of lags used in the Bartlett kernel for estimation of the relevant long-run variance. These are the "traditional asymptotics," as opposed to the "fixed-$b$ asymptotics" which will be discussed in a later Section of this chapter.

### 3.3.1 Asymptotic Results for $\widehat{\eta}_\mu$

The existing results on the limit of $\widehat{\eta}_\mu$ are collected from Kwiatkowski et al. (1992), Shin and Schmidt (1992), Lee and Schmidt (1996), and Lee and Amsler (1997). Table 3.1 shows those results. Note that the case of $d = \frac{1}{2}$ is missing in Table 3.1. Results for $d = \frac{1}{2}$ similar to those in Table 3.1 are not currently available, so this case will be treated separately in this Section. The implications of these results for the asymptotic distribution of $\widehat{\eta}_\mu$ are summarized in Theorem 7.

**Theorem 7.** *Given the data generation process in (2.16), the KPSS statistic, $\widehat{\eta}_\mu$ defined in (2.17) has following asymptotic limits.*

*A. When $\epsilon_t$ is a short-memory process (Kwiatkowski et al. (1992)),*

$$\widehat{\eta}_\mu \Rightarrow \int_0^1 B(r)^2 dr. \tag{2.22}$$

*B. When $\epsilon_t$ is a unit root process (Shin and Schmidt (1992)),*

$$\frac{l}{T}\widehat{\eta}_\mu \Rightarrow \frac{\int_0^1 \left(\int_0^r \underline{W}(s)ds\right)^2 dr}{\int_0^1 \underline{W}(s)^2 ds} \tag{2.23}$$

*implying $\widehat{\eta}_\mu \to \infty$ in probability.*

*C. When $\epsilon_t$ is a fractionally integrated process with $0 < d < 1/2$ (Lee and Schmidt (1996)),*

$$\left(\frac{l}{T}\right)^{2d} \widehat{\eta}_\mu \Rightarrow \int_0^1 B_d(r)^2 dr \tag{2.24}$$

*implying $\widehat{\eta}_\mu \to \infty$ in probability.*

*D. When $\epsilon_t$ is a fractionally integrated process with $1/2 < d < 1$ (Lee and Amsler (1997)),*

$$\left(\frac{l}{T}\right) \widehat{\eta}_\mu \Rightarrow \frac{\int_0^1 \left(\int_0^r \underline{W}_{d_*}(s)ds\right)^2 dr}{\int_0^1 \underline{W}_{d_*}(s)^2 ds} \tag{2.25}$$

*implying $\widehat{\eta}_\mu \to \infty$ in probability.*

The above results cover all of the cases except $d = 1/2$. This case is covered by the following Theorem.

**Theorem 8.** *Given the data generation process in (2.16), the KPSS statistic, $\widehat{\eta}_\mu$ defined in (2.17) diverges to infinity when the error term is an $I(\frac{1}{2})$ process.*

**Proof:** See the Appendix.

Theorems 7 and 8 imply that the KPSS $\widehat{\eta}_\mu$ test is consistent against a unit root and also against $I(d)$ alternatives for all $0 < d < 1$.

### 3.3.2 Asymptotic Results for $\widehat{\eta}_\mu^d$

This Section considers the asymptotic behavior of $\widehat{\eta}_\mu^d$ under stationary short- or long-memory errors, under unit root errors, under nonstationary long-memory errors with $\frac{1}{2} < d < 1$, and under nonstationary long-memory errors with $d = \frac{1}{2}$. Recall

$$\begin{aligned} \widehat{\eta}_\mu^d &\equiv \frac{T^{-2}\sum_{t=2}^{T}\widetilde{S}_t^2}{\widetilde{s}^2(l)}, \text{ with } \widetilde{S}_t \equiv \sum_{j=2}^{t}\Delta y_j = \sum_{j=2}^{t}\Delta\epsilon_j = \epsilon_t - \epsilon_1, \\ \widetilde{s}^2(l) &= \widetilde{\gamma}_0 + 2\sum_{s=1}^{l} w(s,l)\widetilde{\gamma}_s, \text{ and } \widetilde{\gamma}_s = \frac{1}{T-1}\sum_{t=s+2}^{T}\Delta y_t \Delta y_{t-s}, \end{aligned}$$

where $\Delta y_t = y_t - y_{t-1}$, and $\Delta\epsilon_t = \epsilon_t - \epsilon_{t-1}$.

Since $\widehat{\eta}_\mu^d$ is a unit root test statistic, its asymptotic distribution is established under first, the unit root null. Then it will be shown that the limit of the statistic is zero for the case of stationary short memory and for $I(d)$ processes with $0 < d < 1$.

**Proposition 8.** *Under the data generation process in (2.16) with the error term $\epsilon_t$ being a unit root process, and under the assumption that $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$, the statistic, $\widehat{\eta}_\mu^d$ defined in (2.20) weakly converges:*

$$\widehat{\eta}_\mu^d \Rightarrow \int_0^1 W(r)^2 dr, \tag{2.26}$$

*where $W(r)$ is the standard Wiener process.*

**Proof:** See the Appendix.

This is a lower tail test. The 1%, 5% and 10% lower tail critical values are 0.034, 0.056, and 0.076, respectively. These are different from the critical values of the KPSS unit root test in Shin and Schmidt (1992) and Breitung (2002) because the data is differenced instead of demeaning the terms in $\widetilde{S}_t$. As a consequence the result in (2.26) involves an ordinary Wiener process as opposed to a demeaned Wiener process.

The next three results together prove that the limit of the statistic is zero for all cases except the case of unit root errors.

**Theorem 9.** *Under the data generation process in (2.16) with the error term $\epsilon_t$ being a stationary short- or long-memory process and under the assumption that $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$, the statistic, $\widehat{\eta}_\mu^d$ defined in (2.20) has the following limiting distribution:*

$$\left(\frac{l}{T}\right)^{-1} \widehat{\eta}_\mu^d \xrightarrow{d} \frac{\gamma_0 + \epsilon_1^2}{2\gamma_0}, \tag{2.27}$$

*where $\gamma_0 = E\left(\epsilon_t^2\right)$. Therefore, $\widehat{\eta}_\mu^d \xrightarrow{p} 0$.*

**Proof:** See the Appendix.

The next Proposition shows the statistic $\widehat{\eta}_\mu^d$ can distinguish fractionally integrated processes with $\frac{1}{2} < d < 1$ from unit root processes. Recall that this is not the case for the KPSS unit root test (Lee and Amsler (1997)).

**Proposition 9.** *Under the data generation process in (2.16) with the error term $\epsilon_t$ being fractionally integrated with i.i.d. normal innovations and with $\frac{1}{2} < d < 1$ (so, a nonstationary long-memory process), and under the assumption that $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$, the statistic, $\widehat{\eta}_\mu^d$ defined in (2.20) has the following limiting distribution:*

$$\left(\frac{l}{T}\right)^{2d_*} \widehat{\eta}_\mu^d \Rightarrow \int_0^1 W_{d_*}(r)^2 dr, \tag{2.28}$$

*where $d_* = d - 1$ and $W_{d_*}(r)$ is the fractional Brownian motion. Therefore, $\widehat{\eta}_\mu^d \xrightarrow{p} 0$.*

**Proof:** See the Appendix.

Lastly, consider the case of $I\left(\frac{1}{2}\right)$.

**Theorem 10.** *Under the data generation process in (2.16) with the error term $\epsilon_t$ being fractionally integrated with i.i.d. normal innovations and with $d = \frac{1}{2}$ (so, a nonstationary long-memory process) and under the assumption that $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$, the statistic, $\widehat{\eta}_\mu^d$ defined in (2.20) converges to zero.*

**Proof:** See the Appendix.

### 3.3.3 Correct Size and Consistency of the Double-KPSS Test

Now go back to the rejection rule in Section 3.2. The null of integer integration is rejected if the KPSS test rejects short memory and if the unit root test based on $\widehat{\eta}_\mu^d$ rejects unit root. This two-part test has correct size asymptotically and is consistent against $I(d)$ alternatives with $0 < d < 1$.

**Proposition 10.** *Suppose the data generation process is given by (2.16). Under the null hypothesis of integer integration and under the assumption that $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$, the rejection rule*

$$\textit{Reject } H_0 \textit{ if } \widehat{\eta}_\mu > cv^1_{0.95} \textit{ and } \widehat{\eta}_\mu^d < cv^2_{0.05},$$

*gives a test with asymptotic size of 5%, where $cv^1_{0.95}$ is the upper 5% percentile of (2.22) and $cv^2_{0.05}$ is the lower 5% percentile of (2.26). Also, the test is consistent againt the alternative hypothesis of $I(d)$ with $0 < d < 1$.*

**Proof:** The Proposition follows immediately from the results of Sections 3.3.1 and 3.3.2. (1) If the series is $I(0)$, asymptotically the KPSS test will reject with probability 0.05 and the $\widehat{\eta}_\mu^d$ test will reject with probability one, so the Double-KPSS test will reject with probability 0.05. (2) If the series is $I(1)$, the KPSS test will reject with probability one and the $\widehat{\eta}_\mu^d$ test will reject with probability 0.05, so the Double-KPSS test will reject with probability 0.05. (3) If the series is $I(d)$ with $0 < d < 1$, asymptotically both tests will reject with probability one, and so the Double-KPSS test will reject with probability one.

## 3.4 Fixed-$b$ Asymptotic Results

Let $b = \frac{l+1}{T}$, the ratio of the number of lags (plus one) to the sample size. The asymptotic results of the previous Section were obtained under the assumption that $b \to 0$ as $T \to \infty$. This Section discusses the asymptotic distribution of the $\widehat{\eta}_\mu$ and $\widehat{\eta}_\mu^d$ statistics under the "fixed-$b$" assumption that $b$ is held constant as $T \to \infty$. The idea of fixed-$b$ asymptotics was proposed by Kiefer and Vogelsang (2005) and Hashimzade and Vogelsang (2008). The fixed-$b$ approach gives a random limit of the HAC estimator which depends on the choice of kernel and the bandwidth ratio $b$. The fixed-$b$ approach is known to produce a better finite sample approximation to the distribution of test statistics in a variety of settings. Amsler et al. (2009) derived the fixed-$b$ asymptotic distribution of the KPSS $\widehat{\eta}_\mu$ statistic under the $I(0)$ null and under the $I(1)$ alternative. Proposition 11 present their results.

**Proposition 11.** *Given the data generation process in (2.16), under the assumption that $b = \frac{l+1}{T} \in [0,1]$ is held constant as $T$ increases, the KPSS statistic, $\widehat{\eta}_\mu$ defined in (2.17) has the following fixed-b asymptotic limits.*

*When $\epsilon_t$ is a short-memory process,*

$$\widehat{\eta}_\mu \Rightarrow \frac{\int_0^1 B(r)^2 dr}{Q_0(b)}, \tag{2.29}$$

*where* $Q_0(b) \equiv \frac{2}{b}\left[\int_0^1 B(r)^2 dr - \int_0^{1-b} B(r)B(r+b)dr\right]$ *with* $B(r) \equiv W(r) - rW(1)$.

*When $\epsilon_t$ is a unit root process,*

$$\widehat{\eta}_\mu \Rightarrow \frac{\int_0^1 P(r)^2 dr}{Q_1(b)}, \tag{2.30}$$

*where* $Q_1(b) \equiv \frac{2}{b}\left[\int_0^1 P(r)^2 dr - \int_0^{1-b} P(r)P(r+b)dr\right]$ *with* $P(r) \equiv \int_0^r \underline{W}(s)ds$ *with* $\underline{W}(s) \equiv W(s) - \int_0^1 W(u)du$.

**Proof:** See Amsler et al. (2009).

As Proposition 11 shows, the KPSS statistic $\widehat{\eta}_\mu$ has a nondegenerate limit under both $I(0)$ and $I(1)$ data generation processes. So the KPSS $\widehat{\eta}_\mu$ test does not give a consistent test against the unit root under the fixed-$b$ assumption. In the present context, the Double-KPSS test would be conservative (undersized). If the DGP is $I(1)$, the $\widehat{\eta}_\mu^d$ test would reject with probability 0.05, but the KPSS $\widehat{\eta}_\mu$ test would reject with probability less than one, under fixed-$b$ asymptotics. So the probability of both tests rejecting would be less than 0.05. However, these issues should not be regarded as consequential. The assumption of fixed-bandwidth ratio does not recommend any rules for selecting the number of lags. But, no matter how one chooses the number of lags, it will be positive, and the fixed-$b$ critical values will usually give a test of more accurate size than the traditional ($b = 0$) critical values. That is, fixed-$b$ asymptotics is simply viewed as a way of generating a more accurate approximation to the finite sample distribution of the statistic.

The next Proposition provides the fixed-$b$ limits of $\widehat{\eta}_\mu^d$ under $I(0)$ and $I(1)$ data generation processes. To avoid confusion with the previous definition of $b$ and $l$, let the number of lags and the ratio be denoted by $l'$ and $b' = \frac{l'+1}{T-1}$. (We have $T-1$ instead of $T$ because one observation is used up in differencing.)

**Proposition 12.** *Given the data generation process in (2.16), under the assumption that $b' = \frac{l'+1}{T-1} \in [0,1]$ is held constant as $T$ increases, the KPSS statistic, $\widehat{\eta}_\mu^d$ has the following fixed-b asymptotic limits:*

*When $\epsilon_t$ is a short-memory process,*

$$\widehat{\eta}_\mu^d \Rightarrow \frac{\gamma_0 + \epsilon_1^2}{\frac{2}{b'}\gamma_0 + \epsilon_1^2 + \epsilon_\infty^2}, \tag{2.31}$$

*where $\gamma_0 = E\left(\epsilon_t^2\right)$ and $\epsilon_\infty$ denotes the weak limit of $\epsilon_T$ as $T \to \infty$.*

*When $\epsilon_t$ is a unit root process,*

$$\widehat{\eta}_\mu^d \Rightarrow \frac{\int_0^1 W(r)^2 dr}{Q_1^d(b')}, \tag{2.32}$$

*where $W(r)$ is the standard Wiener process and*

$$Q_1^d(b') \equiv \frac{2}{b'}\left[\int_0^1 W(r)^2 dr - \int_0^{1-b'} W(r)W(r+b')dr - \int_{1-b'}^1 W(r)W(1)dr\right] + W(1)^2.$$

**Proof:** See the Appendix.

The fixed-$b$ critical values for $\widehat{\eta}_\mu$ and $\widehat{\eta}_\mu^d$ are simulated using i.i.d. N(0,1) pseudo random numbers with $T = 1,000$ and $50,000$ replications. Table 3.2 provides these fixed-$b$ critical values. The fixed-$b$ critical values for $\widehat{\eta}_\mu$ are slightly different from those in Amsler et al. (2009). This is partly due to randomness of the simulations, but it is also due to a slight difference in the definitions of $b$. They had $b = \frac{l}{T}$ whereas now we have $b = \frac{l+1}{T}$. Since the critical values were simulated using $T = 1,000$, $b = 0.02$ in Amsler et al. (2009) would correspond to $b = 0.021$ in this chapter, for example. For bigger $T$ (i.e. asymptotically) this difference obviously disappears.

The simulation results in Amsler et al. (2009) showed that the size distortion associated with strong short-run persistence of the $I(0)$ DGP can be fixed by using a relatively large number of lags and the corresponding fixed-$b$ critical values. In their results, with the original KPSS critical value being used, as the short run persistence gets higher, the overrejection of the KPSS test gets worse. One can reduce the rejection frequency by using a higher number of lags but now the test becomes subject to an underrejection problem which is translated into low power. However, by taking a relatively large number of lags and using the fixed-$b$ critical values, this problem can be partially fixed. Exactly the same considerations apply to the $\widehat{\eta}_\mu^d$ test and the Double-KPSS test.

## 3.5 Monte Carlo Simulations

This Section reports the results of simulations designed to investigate the finite sample size and power properties of the Double-KPSS test. Some comparisons of the $\widehat{\eta}_\mu^d$ test and the ADF test will be made.

### 3.5.1 Design of the Experiment

The data generating processes to be considered in the simulations are as follows.

1. $I(0)$ DGP:

$$\begin{aligned} y_t &= \mu + \epsilon_t, \\ \epsilon_t &= \rho\epsilon_{t-1} + u_t, \end{aligned} \tag{2.33}$$

with $\mu = 0$, $\epsilon_0 = 0$, $\rho \in \{0, 0.25, 0.5, 0.75, 0.95\}$, $u_t \sim$ i.i.d. $Normal(0, 1)$.

2. $I(1)$ DGP:

$$\begin{aligned} y_t &= \mu + \epsilon_t, \\ \epsilon_t &= \epsilon_{t-1} + \eta_t, \\ \eta_t &= u_t - \phi u_{t-1}, \end{aligned} \tag{2.34}$$

with $\mu = 0$, $\epsilon_0 = u_0 = 0$, $\phi \in \{0, 0.25, 0.5, 0.75, 0.95\}$, $u_t \sim$i.i.d. $Normal(0, 1)$.

3. $I(d)$ DGP:

$$\begin{aligned} y_t &= \mu + \epsilon_t, \\ (1 - L)^d \epsilon_t &= u_t \sim \text{i.i.d. } Normal(0, 1), \end{aligned} \tag{2.35}$$

with $\mu = 0$, and $d \in \{0.1, 0.2, 0.3, 0.4, 0.45, 0.499, 0.5, 0.6, 0.7, 0.75, 0.8, 0.9\}$.

To generate $I(d)$ processes with $0 < d < \frac{1}{2}$ Toeplitz matrix was used (formed from the autocovariances, as in Diebold and Rudebusch (1991)). For the case of $\frac{1}{2} \leq d < 1$, first generated $I(d)$ processes with $-\frac{1}{2} \leq d < 0$ (again using the Toeplitz matrix) and cumulated them to obtain the $I(d)$ processes with $\frac{1}{2} \leq d < 1$. This is the same procedure as in Lee and Schmidt (1996). The experiments considered $T = 50, 100, 200, 500, 1,000$, and

$2,000$ and the number of replications was $5,000$. The numbers of lags used for computing the statistics were $l_0(=0)$, $l_4$, $l_{12}$, $l_{25}$ and $l_{50}$, where

$$l_k \equiv k \cdot \left(\frac{T}{100}\right)^{1/4}. \tag{2.36}$$

For the ADF test, with $p$ lags, $p_4$, $p_{12}$, and $p_{25}$ lags were considered, where $p_k$ is defined in the same way as in equation (2.36).

As a matter of notation, $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ will denote the Double-KPSS test based on $\widehat{\eta}_\mu$ and $\widehat{\eta}_\mu^d$. Similarly $\widehat{\eta}_\mu \times ADF$ will denote the double test but using the ADF test instead of $\widehat{\eta}_\mu^d$.

### 3.5.2 Results with Standard Critical Values

This Section discusses the results for the $\widehat{\eta}_\mu$, $\widehat{\eta}_\mu^d$ and $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ tests, using the "standard" critical values that are valid asymptotically when $l \to \infty$ and $\frac{l}{T} \to 0$ as $T \to \infty$. These results are given in Tables 3.3, 3.4 and 3.5. Each table contains the results for two sample sizes (3.3: $T = 50$ and 100; 3.4: $T = 200$ and 500; 3.5: $T = 1,000$ and $2,000$). The formatting for each sample size is the same.

The results for the KPSS $\widehat{\eta}_\mu$ test are similar to those from previous simulations and will be discussed only briefly. Size under the $I(0)$ null with $\rho = 0$ is essentially correct for $l_0$ but the test is undersized with more lags. The test is oversized when $\rho > 0$ and severely so for the largest values of $\rho$ (like $\rho = 0.95$). Size improves very slowly as $T$ increases. Power against an $I(1)$ alternative rises when $T$ increases, falls as the number of lags increases, and falls as $\phi$ increases (since the series approaches stationarity as $\phi \to 1$). Power against $I(d)$ alternatives grows with $d$, falls as the number of lags increases, and grows (but slowly) as $T$ increases.

The results for the $\widehat{\eta}_\mu^d$ unit root test show a pattern that is similar to what was seen for the KPSS $\widehat{\eta}_\mu$ test, but reversed. Size under the $I(1)$ null with $\phi = 0$ is essentially correct, but the test is undersized with more lags. The test is oversized when $\phi > 0$

and severely so for the largest values of $\phi$ (like $\phi = 0.95$). Size generally improves as $T$ increases. Power against $I(0)$ alternatives rises when $T$ increases, falls as the number of lags increases, and falls as $\rho$ increases (since the series approaches $I(1)$ as $\rho \to 1$). Power against $I(d)$ alternatives grows as $d$ decreases, falls as the number of lags increases, and grows as $T$ increases.

All of these statements would also be true for the ADF test. Some comparisons of the performance of the $\widehat{\eta}_\mu^d$ test and the ADF test will be given in Section 3.5.4.

Now turn to the issue of main interest, the performance of the Double-KPSS ($\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$) test. This test rejects the null of integer integration if both the $\widehat{\eta}_\mu$ short-memory test and the $\widehat{\eta}_\mu^d$ unit root test reject their respective null hypotheses. As a result, the upward size distortions caused by short run dynamics must be smaller for the Double-KPSS test than for either of the individual tests. In many cases the rejection probability for the Double-KPSS test is at least approximately equal to the product of the rejection probabilities for the two component tests, but this is not always the case (The two tests are not independent).

Consider first the size of the Double-KPSS test under the $I(0)$ null. In the most empirically relevant cases, like $l_4$ lags with $T = 100$, or $l_{12}$ lags with $T = 200$ or 500, its size is reasonably accurate, except perhaps for the biggest values of $\rho$. For the largest sample sizes ($T = 1,000$ and $2,000$) the test has fairly accurate size, except for the case of $\rho = 0.95$, if the test uses $l_{12} \times l_{12}$ or $l_{25} \times l_{25}$ lags. However, the size of the test does not improve uniformly as $T$ increases, because loosely speaking the power of the unit root test goes to one faster than the size of the short-memory test goes to 0.05. But as a general statement the size of the test is surprisingly good over a broad range of values of $\rho$.

Now consider the size of the test under the $I(1)$ null. Once again the size is reasonably accurate, except perhaps for the biggest values of $\phi$, if reasonable numbers of lags, like $l_4$ lags with $T = 100$, or $l_{12}$ lags with $T = 200$ or 500 are used. The test is if anything undersized (due to the use of the standard critical values despite the positive number of lags) for the smaller values of $\phi$. For the largest values of $T$ ($1,000$ and $2,000$), size is quite good with $l_{12} \times l_{12}$ or $l_{25} \times l_{25}$ lags, except when $\phi = 0.75$ or 0.95.

Finally, consider the power of the test against $I(d)$ alternatives. Now there is a potential problem, because if one uses the numbers of lags mentioned above as sufficient to control size, power is low. For example, if $l_{12}$ lags, with $T = 200$ is used, the highest power is only 0.099 (against $d = 0.4$) and with $T = 500$ the highest power is 0.386 (against $d = 0.5$). Of course, there is a trade-off between size and power. If one uses only $l_4$ lags, maximal power is 0.488 for $T = 200$ and 0.786 for $T = 500$. But with only $l_4$ lags, there are large size distortions under the null for the larger values of $\rho$ (for the $I(0)$ null) or $\phi$ (for the $I(1)$ null). It takes a very large sample size (like $T = 1,000$ or $2,000$) to have reasonable power with $l_{12}$ lags.

So, what can one conclude from these simulations? In a view the main practical question is how large the sample size needs to be so that one can reasonably conclude that a rejection from the test is due to its power against a fractional alternative, as opposed to size distortions of one or both of the two component tests. This obviously will depend on the values of $d$ against which we require power, as well as the values of the nuisance parameters that we want the null hypothesis to encompass. As an extreme example, there is no hope of success if we want to include in the $I(0)$ null $AR$ processes with local to unity roots, or if we want to include in the $I(1)$ null $ARIMA(0,1,1)$ processes with local to unity $MA$ roots.

The simulations results seem to indicate that the test can in fact reasonably distinguish fractional integration from non-extreme $I(0)$ or $I(1)$ processes, but that it will take a large sample size to do so. For example, for $T = 500$ and for the tests using $l_{12}$ lags, power for $d$ in the range $[0.3, 0.7]$ is at least twice as large as the maximal size distortion for $I(0)$ processes with $AR$ roots less than or equal to 0.75 or for $I(1)$ processes with $MA$ roots less than or equal to 0.75. For smaller sample sizes, this statement would not be true, and to make a similar statement that is true would require a smaller range of $d$ and/or a more restrictive range of $AR$ or $MA$ roots. Conversely, to make a similar statement that is true for a larger range of $d$ or of $AR$ and $MA$ parameters will require a larger sample size. For example, with $T = 2000$, power for $d$ in the range $[0.2, 0.8]$ is at least twice as large as the maximal size distortion for $I(0)$ processes with $\rho$ less than or equal to 0.75 or for $I(1)$ processes with $\phi$ less than or equal to 0.75. The obvious problem with these statements is that, for economic time series data, $T = 2000$, or for that matter $T = 500$, is a very large sample size.

### 3.5.3 Results with Fixed-$b$ Critical Values

The fixed-$b$ critical values, for the relevant values of $b$, are smaller than the traditional critical values for the KPSS $\widehat{\eta}_\mu$ test (an upper tail test) and larger for the $\widehat{\eta}_\mu^d$ test (a lower tail test). The fixed-$b$ critical values will therefore lead to more rejections than the traditional critical values, if the same number of lags is used in both cases. If the number of lags increases with sample size but more slowly than sample size, the difference in the critical values (and the rejection probabilities) will go to zero since $b$ will go to zero.

The fixed-$b$ critical values are very successful in removing the underrejection problem that occurs for the KPSS $\widehat{\eta}_\mu$ and $\widehat{\eta}_\mu^d$ tests when there are no short-run dynamics and the sample size is not large. For example, for KPSS $\widehat{\eta}_\mu$ with $T = 50$ and $I(0)$ data with $\rho = 0$, and with $l_{12}$ lags, compare size of 0.014 with traditional critical values to 0.053 with fixed-$b$ critical values. Or, for the $\widehat{\eta}_\mu^d$ test with $T = 50$ and $I(1)$ data with $\phi = 0$, and with $l_{12}$ lags, compare 0.000 with traditional critical values to 0.039 with fixed-$b$ critical values. For the Double-KPSS test there is also improvement in size in these cases from using the fixed-$b$ critical values, but the improvement is not so striking.

Upward size distortions in the presence of short run dynamics ($I(0)$ data with large positive $\rho$ or $I(1)$ data with large positive $\phi$) are worse when the fixed-$b$ critical values are used. Also power against $I(d)$ alternatives is higher when the fixed-$b$ critical values are used. However, these differences are not large when the sample size is big enough for us to have reasonable power (e.g., $T$ greater than or equal to 500). Of course, that is because the rule for the choice of lags in the simulations implies that $b$ goes to zero as $T$ grows, but that is a reasonable feature for such a rule to have. Using the fixed-$b$ critical values is recommended but recognize that if the number of lags is chosen reasonably this is likely to not make much difference.

### 3.5.4 Comparison of the $\widehat{\eta}_{\mu}^{d}$ Test and the ADF Test

Although it is not the focus of this research, a new unit root test has been proposed in this chapter and it is relevant to ask how it compares to other existing unit root tests. There are of course a great many other existing unit root tests. The ADF $\widehat{\tau}_{\mu}$ test will be taken as a standard of comparison . The number of lagged differences included in the ADF regression is denoted as $p$, and in making comparisons of size and power a value of $p$ will be matched to the same value of $l$, the number of lags used for long run variance estimation in the $\widehat{\eta}_{\mu}^{d}$ test.

Table 3.9 gives size and power for the ADF test for $T = 50$, 100, 200 and 500, and these results can be compared to the results previously given in Table 3.3 and 3.4.

In terms of the size of the test, the results are mixed. However, for the larger sample sizes, the ADF test with $p_{12}$ lags has smaller size distortions than the $\widehat{\eta}_{\mu}^{d}$ test with $l_{12}$ lags for the larger values of $\phi$. The ADF test generally has higher power against $I(0)$ alternatives, while the $\widehat{\eta}_{\mu}^{d}$ test has higher power against $I(d)$ alternatives except when $d$ is very small (e.g. 0.1).

If one compares the Double-KPSS ($\widehat{\eta}_{\mu} \times \widehat{\eta}_{\mu}^{d}$) test to the $\widehat{\eta}_{\mu} \times ADF$ test, similar statements apply, but the differences are much smaller. In fact, the similarities between these two double tests far outweigh the differences. Unsurprisingly, perhaps, the precise choice of unit root test to use is not the main issue here.

## 3.6 Conclusions

This chapter proposed a Double KPSS test to test the null of integer integration ($I(0)$ or $I(1)$) against the alternative of fractional integration ($I(d)$ with $d$ between zero and one). The null of integer integration is rejected if the KPSS test rejects the null of short memory and a unit root test rejects the null of a unit root. A new unit root test was suggested for use in this testing procedure, but any other unit root test like the ADF test is also possible. This would be a good preliminary test to use before estimating a fractional model. An alternative, of course, is to just estimate the fractional model and see whether the estimated $d$ is significantly different from zero and from one. However, there appears to be no clear consensus in the existing literature on how to allow for short-run dynamics in estimating $d$ and conducting inference about it.

The consistency of the test were proved. The main practical question is how large the sample size needs to be so that one can reasonably conclude that a rejection from the test is due to its power against a fractional alternative, as opposed to size distortions of the two component tests. The simulations results seem to indicate that the test can in fact distinguish fractional integration from non-extreme $I(0)$ or $I(1)$ processes, but that it will take a very large sample size to do reliably. This is not a surprising result. It takes a lot of data to distinguish $I(0)$ from $I(1)$ processes, if the range of short-run dynamics is not severely restricted. Now we are trying to do more, for example, to distinguish a unit root process from an $I(d)$ process with $d = 0.8$. An important contribution of this chapter is to try to quantify how much data that takes. The required sample sizes would be very large indeed for macroeconomic applications, but perhaps not for applications in finance.

Table 3.1: Summary of the existing asymptotic results for $\widehat{\eta}_\mu$

| $\epsilon_t \sim$ | $H_0$: $I(0)$ | $H_0$: $I(1)$ |
|---|---|---|
| $S_t = \sum_{j=1}^t e_j$ : | $\frac{1}{\sqrt{T}} S_{[rT]} \Rightarrow \sigma B(r)$, where $B(r) = W(r) - rW(1)$ | $\frac{1}{T^{3/2}} S_{[rT]} \Rightarrow \sigma \int_0^r \underline{W}(s)ds$, where $\underline{W}(s) = W(s) - \int_0^1 W(u)du$ |
| $\sum S_t^2$ : | $\frac{1}{T^2} \sum_{t=1}^T S_t^2 \Rightarrow \sigma^2 \int_0^1 B(r)^2 dr$ | $\frac{1}{T^4} \sum_{t=1}^T S_t^2 \Rightarrow \sigma^2 \int_0^1 \left( \int_0^r \underline{W}(s)ds \right)^2 dr$ |
| $s^2(l)$ : | $s^2(l) \xrightarrow{p} \sigma^2$ | $\frac{1}{lT} s^2(l) \Rightarrow \sigma^2 \int_0^1 \underline{W}(s)^2 ds$ |
| $\epsilon_t \sim$ | $H_1$: $I(d), 0 < d < 1/2$ | $H_1$: $I(d), 1/2 < d < 1$ |
| $S_t = \sum_{j=1}^t e_j$ : | $\frac{1}{T^{d+1/2}} S_{[rT]} \Rightarrow \omega_d B_d(r)$, where $B_d(r) = W_d(r) - rW_d(1)$ | $\frac{1}{T^{d_*+3/2}} S_{[rT]} \Rightarrow \omega_{d_*} \int_0^r \underline{W}_{d_*}(s)ds$, where $d_* = d - 1$ and $\underline{W}_{d_*}(r) = W_{d_*}(r) - \int_0^1 W_{d_*}(s)ds$ |
| $\sum S_t^2$ : | $\frac{1}{T^{2(d+1)}} \sum_{t=1}^T S_t^2 \Rightarrow \omega_d^2 \int_0^1 B_d(r)^2 dr$ | $\frac{1}{T^{2d_*+4}} \sum_{t=1}^T S_t^2 \Rightarrow \omega_{d_*}^2 \int_0^1 \left( \int_0^r \underline{W}_{d_*}(s)ds \right)^2 dr$ |
| $s^2(l)$ : | $l^{-2d} s^2(l) \xrightarrow{p} \omega_d^2$ | $\frac{1}{lT^{2d_*+1}} s^2(l) \Rightarrow \omega_{d_*}^2 \int_0^1 \underline{W}_{d_*}(s)^2 ds$ |

Table 3.2: Fixed-$b$ Critical Values for $\widehat{\eta}_{\mu}$ and $\widehat{\eta}_{\mu}^{d}$, Bartlett kernel, $l = bT - 1$, $l' = b'(T-1) - 1$

| | $\widehat{\eta}_{\mu}$ upper tail | | $\widehat{\eta}_{\mu}^{d}$ lower tail | | | $\widehat{\eta}_{\mu}$ upper tail | | $\widehat{\eta}_{\mu}^{d}$ lower tail | |
|---|---|---|---|---|---|---|---|---|---|
| $b$ | 5% | 1% | 1% | 5% | $b$ | 5% | 1% | 1% | 5% |
| 0 | 0.463 | 0.739 | 0.034 | 0.056 | 0.52 | 0.405 | 0.445 | 0.100 | 0.144 |
| 0.02 | 0.453 | 0.705 | 0.038 | 0.060 | 0.54 | 0.410 | 0.453 | 0.101 | 0.145 |
| 0.04 | 0.446 | 0.673 | 0.042 | 0.063 | 0.56 | 0.414 | 0.461 | 0.101 | 0.147 |
| 0.06 | 0.439 | 0.639 | 0.045 | 0.067 | 0.58 | 0.419 | 0.468 | 0.102 | 0.148 |
| 0.08 | 0.434 | 0.609 | 0.049 | 0.071 | 0.60 | 0.424 | 0.473 | 0.102 | 0.150 |
| 0.10 | 0.428 | 0.582 | 0.053 | 0.075 | 0.62 | 0.429 | 0.479 | 0.103 | 0.151 |
| 0.12 | 0.421 | 0.561 | 0.057 | 0.078 | 0.64 | 0.432 | 0.485 | 0.104 | 0.151 |
| 0.14 | 0.416 | 0.541 | 0.062 | 0.082 | 0.66 | 0.436 | 0.490 | 0.104 | 0.152 |
| 0.16 | 0.409 | 0.522 | 0.066 | 0.086 | 0.68 | 0.439 | 0.493 | 0.104 | 0.153 |
| 0.18 | 0.403 | 0.504 | 0.069 | 0.090 | 0.70 | 0.442 | 0.496 | 0.105 | 0.154 |
| 0.20 | 0.398 | 0.489 | 0.073 | 0.095 | 0.72 | 0.445 | 0.499 | 0.106 | 0.155 |
| 0.22 | 0.394 | 0.476 | 0.077 | 0.099 | 0.74 | 0.448 | 0.499 | 0.106 | 0.156 |
| 0.24 | 0.391 | 0.464 | 0.080 | 0.103 | 0.76 | 0.449 | 0.501 | 0.106 | 0.157 |
| 0.26 | 0.388 | 0.455 | 0.083 | 0.107 | 0.78 | 0.452 | 0.500 | 0.107 | 0.158 |
| 0.28 | 0.385 | 0.447 | 0.086 | 0.111 | 0.80 | 0.454 | 0.498 | 0.107 | 0.159 |
| 0.30 | 0.383 | 0.441 | 0.088 | 0.115 | 0.82 | 0.456 | 0.498 | 0.107 | 0.160 |
| 0.32 | 0.381 | 0.435 | 0.090 | 0.119 | 0.84 | 0.458 | 0.496 | 0.108 | 0.160 |
| 0.34 | 0.380 | 0.430 | 0.092 | 0.122 | 0.86 | 0.462 | 0.494 | 0.108 | 0.161 |
| 0.36 | 0.381 | 0.426 | 0.093 | 0.126 | 0.88 | 0.465 | 0.490 | 0.108 | 0.162 |
| 0.38 | 0.380 | 0.424 | 0.095 | 0.129 | 0.90 | 0.468 | 0.488 | 0.109 | 0.163 |
| 0.40 | 0.381 | 0.422 | 0.095 | 0.132 | 0.92 | 0.473 | 0.487 | 0.109 | 0.163 |
| 0.42 | 0.383 | 0.421 | 0.096 | 0.135 | 0.94 | 0.478 | 0.487 | 0.109 | 0.164 |
| 0.44 | 0.386 | 0.422 | 0.098 | 0.137 | 0.96 | 0.484 | 0.488 | 0.110 | 0.165 |
| 0.46 | 0.390 | 0.425 | 0.098 | 0.139 | 0.98 | 0.491 | 0.492 | 0.110 | 0.166 |
| 0.48 | 0.394 | 0.430 | 0.099 | 0.141 | 1 | NA | NA | 0.110 | 0.166 |
| 0.50 | 0.400 | 0.437 | 0.100 | 0.142 | | | | | |

Table 3.3: Size and Power Using Standard 5% Critical Values with Traditional Lag Choices, $T =$ 50 and 100

| T=50 | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}_\mu^d$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_0$ | $l_4$ | $l_{12}$ | $l_0$ | $l_4$ | $l_{12}$ | $l_0 \times l_0$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.053 | 0.045 | 0.014 | 0.959 | 0.478 | 0.001 | 0.050 | 0.015 | 0.000 |
| 0.25 | 0.149 | 0.062 | 0.016 | 0.923 | 0.484 | 0.001 | 0.134 | 0.021 | 0.000 |
| 0.50 | 0.336 | 0.102 | 0.021 | 0.867 | 0.478 | 0.001 | 0.280 | 0.031 | 0.000 |
| 0.75 | 0.653 | 0.214 | 0.035 | 0.641 | 0.324 | 0.001 | 0.379 | 0.030 | 0.000 |
| 0.95 | 0.924 | 0.530 | 0.118 | 0.108 | 0.048 | 0.001 | 0.082 | 0.003 | 0.000 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi$= 0 | 0.961 | 0.711 | 0.353 | 0.042 | 0.021 | 0.000 | 0.030 | 0.002 | 0.000 |
| 0.25 | 0.944 | 0.703 | 0.348 | 0.141 | 0.028 | 0.000 | 0.112 | 0.003 | 0.000 |
| 0.50 | 0.895 | 0.679 | 0.331 | 0.400 | 0.064 | 0.000 | 0.321 | 0.011 | 0.000 |
| 0.75 | 0.709 | 0.571 | 0.262 | 0.786 | 0.208 | 0.000 | 0.518 | 0.062 | 0.000 |
| 0.95 | 0.126 | 0.100 | 0.035 | 0.955 | 0.466 | 0.001 | 0.117 | 0.036 | 0.000 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.133 | 0.082 | 0.021 | 0.946 | 0.464 | 0.001 | 0.124 | 0.029 | 0.000 |
| 0.2 | 0.252 | 0.138 | 0.036 | 0.925 | 0.430 | 0.001 | 0.227 | 0.043 | 0.000 |
| 0.3 | 0.399 | 0.198 | 0.053 | 0.890 | 0.379 | 0.001 | 0.340 | 0.045 | 0.000 |
| 0.4 | 0.546 | 0.276 | 0.076 | 0.820 | 0.322 | 0.000 | 0.427 | 0.050 | 0.000 |
| 0.45 | 0.616 | 0.322 | 0.084 | 0.774 | 0.282 | 0.000 | 0.450 | 0.047 | 0.000 |
| 0.499 | 0.674 | 0.372 | 0.100 | 0.723 | 0.251 | 0.001 | 0.458 | 0.042 | 0.000 |
| 0.5 | 0.679 | 0.369 | 0.100 | 0.720 | 0.239 | 0.000 | 0.457 | 0.043 | 0.000 |
| 0.6 | 0.775 | 0.456 | 0.131 | 0.572 | 0.167 | 0.000 | 0.402 | 0.027 | 0.000 |
| 0.7 | 0.851 | 0.533 | 0.174 | 0.389 | 0.110 | 0.000 | 0.290 | 0.017 | 0.000 |
| 0.75 | 0.879 | 0.571 | 0.200 | 0.294 | 0.085 | 0.000 | 0.222 | 0.012 | 0.000 |
| 0.8 | 0.907 | 0.603 | 0.229 | 0.215 | 0.066 | 0.000 | 0.164 | 0.009 | 0.000 |
| 0.9 | 0.939 | 0.662 | 0.283 | 0.102 | 0.034 | 0.000 | 0.077 | 0.003 | 0.000 |

| T=100 | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}_\mu^d$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_0$ | $l_4$ | $l_{12}$ | $l_0$ | $l_4$ | $l_{12}$ | $l_0 \times l_0$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.043 | 0.040 | 0.027 | 0.999 | 0.751 | 0.040 | 0.043 | 0.027 | 0.001 |
| 0.25 | 0.143 | 0.052 | 0.033 | 0.993 | 0.764 | 0.046 | 0.142 | 0.036 | 0.001 |
| 0.50 | 0.361 | 0.088 | 0.041 | 0.983 | 0.797 | 0.064 | 0.354 | 0.063 | 0.000 |
| 0.75 | 0.718 | 0.194 | 0.068 | 0.944 | 0.784 | 0.100 | 0.673 | 0.137 | 0.001 |
| 0.95 | 0.978 | 0.591 | 0.250 | 0.262 | 0.172 | 0.029 | 0.247 | 0.055 | 0.000 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.993 | 0.821 | 0.578 | 0.046 | 0.029 | 0.006 | 0.044 | 0.009 | 0.000 |
| 0.25 | 0.989 | 0.818 | 0.577 | 0.157 | 0.041 | 0.006 | 0.149 | 0.015 | 0.000 |
| 0.50 | 0.978 | 0.809 | 0.571 | 0.451 | 0.091 | 0.006 | 0.431 | 0.040 | 0.000 |
| 0.75 | 0.905 | 0.747 | 0.541 | 0.863 | 0.332 | 0.008 | 0.769 | 0.188 | 0.001 |
| 0.95 | 0.300 | 0.265 | 0.186 | 0.998 | 0.718 | 0.032 | 0.299 | 0.178 | 0.002 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.156 | 0.091 | 0.052 | 0.997 | 0.738 | 0.041 | 0.155 | 0.061 | 0.001 |
| 0.2 | 0.346 | 0.173 | 0.086 | 0.992 | 0.713 | 0.037 | 0.343 | 0.113 | 0.001 |
| 0.3 | 0.539 | 0.272 | 0.131 | 0.983 | 0.670 | 0.034 | 0.528 | 0.164 | 0.001 |
| 0.4 | 0.715 | 0.373 | 0.181 | 0.958 | 0.609 | 0.031 | 0.681 | 0.201 | 0.000 |
| 0.45 | 0.783 | 0.424 | 0.214 | 0.936 | 0.564 | 0.026 | 0.726 | 0.207 | 0.001 |
| 0.499 | 0.830 | 0.473 | 0.238 | 0.907 | 0.509 | 0.029 | 0.744 | 0.202 | 0.000 |
| 0.5 | 0.830 | 0.474 | 0.241 | 0.904 | 0.509 | 0.028 | 0.742 | 0.201 | 0.000 |
| 0.6 | 0.908 | 0.561 | 0.310 | 0.775 | 0.370 | 0.023 | 0.690 | 0.151 | 0.000 |
| 0.7 | 0.954 | 0.640 | 0.380 | 0.572 | 0.235 | 0.018 | 0.533 | 0.091 | 0.000 |
| 0.75 | 0.967 | 0.678 | 0.410 | 0.454 | 0.175 | 0.015 | 0.427 | 0.069 | 0.000 |
| 0.8 | 0.975 | 0.710 | 0.440 | 0.330 | 0.128 | 0.013 | 0.311 | 0.047 | 0.000 |
| 0.9 | 0.985 | 0.774 | 0.513 | 0.139 | 0.062 | 0.010 | 0.130 | 0.022 | 0.000 |

Table 3.4: Size and Power Using Standard 5% Critical Values with Traditional Lag Choices, $T =$ 200 and 500

| **T=200** | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}^d_\mu$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}^d_\mu$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.044 | 0.040 | 0.029 | 0.945 | 0.555 | 0.008 | 0.041 | 0.022 | 0.000 |
| 0.25 | 0.061 | 0.047 | 0.031 | 0.951 | 0.569 | 0.009 | 0.057 | 0.026 | 0.000 |
| 0.50 | 0.098 | 0.055 | 0.033 | 0.965 | 0.615 | 0.012 | 0.093 | 0.033 | 0.001 |
| 0.75 | 0.223 | 0.079 | 0.042 | 0.975 | 0.706 | 0.021 | 0.215 | 0.050 | 0.000 |
| 0.95 | 0.711 | 0.311 | 0.127 | 0.574 | 0.333 | 0.032 | 0.358 | 0.046 | 0.000 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.948 | 0.720 | 0.524 | 0.037 | 0.022 | 0.004 | 0.026 | 0.004 | 0.000 |
| 0.25 | 0.948 | 0.721 | 0.525 | 0.058 | 0.023 | 0.003 | 0.041 | 0.004 | 0.000 |
| 0.50 | 0.945 | 0.720 | 0.523 | 0.127 | 0.031 | 0.004 | 0.100 | 0.006 | 0.000 |
| 0.75 | 0.919 | 0.707 | 0.514 | 0.445 | 0.094 | 0.003 | 0.385 | 0.028 | 0.000 |
| 0.95 | 0.557 | 0.466 | 0.340 | 0.917 | 0.499 | 0.006 | 0.501 | 0.211 | 0.001 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.132 | 0.084 | 0.051 | 0.940 | 0.550 | 0.009 | 0.121 | 0.046 | 0.001 |
| 0.2 | 0.264 | 0.143 | 0.082 | 0.926 | 0.534 | 0.008 | 0.241 | 0.074 | 0.001 |
| 0.3 | 0.403 | 0.219 | 0.117 | 0.904 | 0.500 | 0.008 | 0.358 | 0.096 | 0.001 |
| 0.4 | 0.540 | 0.292 | 0.161 | 0.863 | 0.439 | 0.008 | 0.457 | 0.099 | 0.000 |
| 0.45 | 0.597 | 0.334 | 0.183 | 0.831 | 0.391 | 0.009 | 0.483 | 0.094 | 0.000 |
| 0.499 | 0.646 | 0.379 | 0.214 | 0.785 | 0.348 | 0.008 | 0.486 | 0.084 | 0.000 |
| 0.5 | 0.647 | 0.381 | 0.215 | 0.785 | 0.350 | 0.007 | 0.488 | 0.091 | 0.000 |
| 0.6 | 0.742 | 0.457 | 0.271 | 0.646 | 0.237 | 0.007 | 0.447 | 0.057 | 0.000 |
| 0.7 | 0.819 | 0.526 | 0.333 | 0.442 | 0.145 | 0.006 | 0.321 | 0.029 | 0.000 |
| 0.75 | 0.852 | 0.558 | 0.364 | 0.341 | 0.111 | 0.006 | 0.253 | 0.022 | 0.000 |
| 0.8 | 0.879 | 0.594 | 0.391 | 0.246 | 0.079 | 0.006 | 0.184 | 0.016 | 0.000 |
| 0.9 | 0.923 | 0.658 | 0.457 | 0.105 | 0.042 | 0.005 | 0.074 | 0.006 | 0.000 |

| **T=500** | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}^d_\mu$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}^d_\mu$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.054 | 0.049 | 0.043 | 0.997 | 0.873 | 0.556 | 0.053 | 0.042 | 0.023 |
| 0.25 | 0.067 | 0.054 | 0.044 | 0.997 | 0.883 | 0.570 | 0.067 | 0.046 | 0.025 |
| 0.50 | 0.094 | 0.060 | 0.047 | 0.999 | 0.918 | 0.616 | 0.094 | 0.054 | 0.028 |
| 0.75 | 0.192 | 0.084 | 0.056 | 1.000 | 0.971 | 0.728 | 0.192 | 0.080 | 0.039 |
| 0.95 | 0.728 | 0.322 | 0.143 | 0.995 | 0.969 | 0.784 | 0.724 | 0.305 | 0.081 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.992 | 0.901 | 0.730 | 0.045 | 0.038 | 0.023 | 0.041 | 0.020 | 0.003 |
| 0.25 | 0.992 | 0.901 | 0.729 | 0.059 | 0.042 | 0.025 | 0.054 | 0.022 | 0.003 |
| 0.50 | 0.992 | 0.901 | 0.728 | 0.123 | 0.054 | 0.027 | 0.117 | 0.030 | 0.004 |
| 0.75 | 0.990 | 0.897 | 0.728 | 0.440 | 0.160 | 0.046 | 0.431 | 0.114 | 0.012 |
| 0.95 | 0.888 | 0.791 | 0.661 | 0.981 | 0.778 | 0.400 | 0.870 | 0.601 | 0.231 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.173 | 0.119 | 0.084 | 0.996 | 0.868 | 0.555 | 0.173 | 0.101 | 0.047 |
| 0.2 | 0.365 | 0.225 | 0.145 | 0.995 | 0.860 | 0.546 | 0.362 | 0.192 | 0.075 |
| 0.3 | 0.555 | 0.341 | 0.220 | 0.991 | 0.839 | 0.515 | 0.549 | 0.281 | 0.101 |
| 0.4 | 0.721 | 0.461 | 0.297 | 0.980 | 0.799 | 0.460 | 0.705 | 0.359 | 0.113 |
| 0.45 | 0.785 | 0.515 | 0.339 | 0.967 | 0.760 | 0.418 | 0.755 | 0.380 | 0.109 |
| 0.499 | 0.832 | 0.566 | 0.379 | 0.950 | 0.715 | 0.363 | 0.786 | 0.384 | 0.095 |
| 0.5 | 0.830 | 0.568 | 0.381 | 0.947 | 0.715 | 0.364 | 0.782 | 0.386 | 0.099 |
| 0.6 | 0.906 | 0.662 | 0.461 | 0.844 | 0.565 | 0.242 | 0.756 | 0.339 | 0.063 |
| 0.7 | 0.950 | 0.746 | 0.530 | 0.642 | 0.361 | 0.139 | 0.600 | 0.221 | 0.029 |
| 0.75 | 0.964 | 0.778 | 0.571 | 0.507 | 0.272 | 0.102 | 0.477 | 0.163 | 0.020 |
| 0.8 | 0.974 | 0.808 | 0.609 | 0.366 | 0.186 | 0.076 | 0.345 | 0.108 | 0.014 |
| 0.9 | 0.987 | 0.861 | 0.672 | 0.141 | 0.083 | 0.042 | 0.134 | 0.045 | 0.008 |

Table 3.5: Size and Power Using Standard 5% Critical Values with Traditional Lag Choices, $T =$ 1,000 and 2,000

| **T=1,000** | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}_\mu^d$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.047 | 0.045 | 0.042 | 1.000 | 0.966 | 0.804 | 0.047 | 0.043 | 0.034 |
| 0.25 | 0.057 | 0.049 | 0.044 | 1.000 | 0.971 | 0.817 | 0.057 | 0.047 | 0.037 |
| 0.50 | 0.076 | 0.056 | 0.046 | 1.000 | 0.985 | 0.866 | 0.076 | 0.055 | 0.040 |
| 0.75 | 0.147 | 0.073 | 0.055 | 1.000 | 0.999 | 0.946 | 0.147 | 0.073 | 0.052 |
| 0.95 | 0.641 | 0.272 | 0.133 | 1.000 | 1.000 | 0.996 | 0.641 | 0.272 | 0.132 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.998 | 0.959 | 0.849 | 0.049 | 0.042 | 0.034 | 0.048 | 0.030 | 0.013 |
| 0.25 | 0.998 | 0.959 | 0.848 | 0.061 | 0.046 | 0.037 | 0.060 | 0.034 | 0.014 |
| 0.50 | 0.998 | 0.959 | 0.848 | 0.117 | 0.064 | 0.042 | 0.116 | 0.050 | 0.016 |
| 0.75 | 0.997 | 0.958 | 0.848 | 0.390 | 0.174 | 0.085 | 0.387 | 0.148 | 0.041 |
| 0.95 | 0.975 | 0.931 | 0.819 | 0.988 | 0.851 | 0.599 | 0.963 | 0.786 | 0.467 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.197 | 0.139 | 0.106 | 1.000 | 0.964 | 0.802 | 0.197 | 0.133 | 0.085 |
| 0.2 | 0.414 | 0.273 | 0.192 | 1.000 | 0.962 | 0.795 | 0.414 | 0.262 | 0.150 |
| 0.3 | 0.630 | 0.422 | 0.299 | 0.999 | 0.950 | 0.780 | 0.630 | 0.398 | 0.227 |
| 0.4 | 0.799 | 0.576 | 0.406 | 0.997 | 0.923 | 0.748 | 0.797 | 0.524 | 0.293 |
| 0.45 | 0.860 | 0.637 | 0.459 | 0.994 | 0.903 | 0.706 | 0.853 | 0.567 | 0.311 |
| 0.499 | 0.899 | 0.695 | 0.507 | 0.986 | 0.866 | 0.652 | 0.886 | 0.591 | 0.305 |
| 0.5 | 0.899 | 0.696 | 0.509 | 0.984 | 0.870 | 0.645 | 0.884 | 0.593 | 0.300 |
| 0.6 | 0.953 | 0.792 | 0.600 | 0.925 | 0.729 | 0.504 | 0.879 | 0.555 | 0.257 |
| 0.7 | 0.978 | 0.859 | 0.681 | 0.740 | 0.513 | 0.322 | 0.718 | 0.410 | 0.159 |
| 0.75 | 0.987 | 0.885 | 0.713 | 0.603 | 0.395 | 0.239 | 0.591 | 0.319 | 0.112 |
| 0.8 | 0.991 | 0.906 | 0.747 | 0.446 | 0.291 | 0.170 | 0.438 | 0.234 | 0.076 |
| 0.9 | 0.995 | 0.939 | 0.800 | 0.186 | 0.126 | 0.081 | 0.182 | 0.099 | 0.032 |

| **T=2,000** | $\widehat{\eta}_\mu$ | | | $\widehat{\eta}_\mu^d$ | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 0.048 | 0.046 | 0.044 | 1.000 | 0.995 | 0.940 | 0.048 | 0.046 | 0.041 |
| 0.25 | 0.058 | 0.050 | 0.045 | 1.000 | 0.996 | 0.949 | 0.058 | 0.050 | 0.043 |
| 0.50 | 0.074 | 0.054 | 0.047 | 1.000 | 0.999 | 0.971 | 0.074 | 0.054 | 0.046 |
| 0.75 | 0.132 | 0.067 | 0.054 | 1.000 | 1.000 | 0.994 | 0.132 | 0.067 | 0.054 |
| 0.95 | 0.599 | 0.242 | 0.122 | 1.000 | 1.000 | 1.000 | 0.599 | 0.242 | 0.122 |
| $\epsilon_t \sim I(1)$ | | (Power) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 1.000 | 0.989 | 0.940 | 0.051 | 0.045 | 0.040 | 0.051 | 0.042 | 0.027 |
| 0.25 | 1.000 | 0.989 | 0.940 | 0.062 | 0.049 | 0.041 | 0.062 | 0.046 | 0.028 |
| 0.50 | 1.000 | 0.989 | 0.941 | 0.111 | 0.066 | 0.048 | 0.110 | 0.062 | 0.033 |
| 0.75 | 1.000 | 0.989 | 0.941 | 0.375 | 0.165 | 0.092 | 0.375 | 0.159 | 0.069 |
| 0.95 | 0.998 | 0.982 | 0.927 | 0.995 | 0.878 | 0.677 | 0.993 | 0.861 | 0.614 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 0.223 | 0.163 | 0.125 | 1.000 | 0.995 | 0.938 | 0.223 | 0.163 | 0.117 |
| 0.2 | 0.496 | 0.345 | 0.254 | 1.000 | 0.994 | 0.937 | 0.496 | 0.344 | 0.237 |
| 0.3 | 0.745 | 0.533 | 0.397 | 1.000 | 0.993 | 0.925 | 0.745 | 0.529 | 0.367 |
| 0.4 | 0.890 | 0.693 | 0.523 | 1.000 | 0.985 | 0.895 | 0.890 | 0.683 | 0.467 |
| 0.45 | 0.932 | 0.764 | 0.588 | 1.000 | 0.976 | 0.869 | 0.932 | 0.745 | 0.505 |
| 0.499 | 0.958 | 0.815 | 0.651 | 0.999 | 0.957 | 0.834 | 0.957 | 0.777 | 0.528 |
| 0.5 | 0.959 | 0.817 | 0.651 | 0.998 | 0.955 | 0.829 | 0.958 | 0.779 | 0.525 |
| 0.6 | 0.987 | 0.892 | 0.753 | 0.977 | 0.858 | 0.692 | 0.965 | 0.755 | 0.493 |
| 0.7 | 0.995 | 0.938 | 0.820 | 0.849 | 0.651 | 0.489 | 0.844 | 0.596 | 0.361 |
| 0.75 | 0.996 | 0.952 | 0.848 | 0.707 | 0.527 | 0.369 | 0.704 | 0.487 | 0.271 |
| 0.8 | 0.998 | 0.966 | 0.873 | 0.552 | 0.385 | 0.258 | 0.550 | 0.359 | 0.185 |
| 0.9 | 1.000 | 0.981 | 0.910 | 0.217 | 0.157 | 0.115 | 0.216 | 0.146 | 0.080 |

Table 3.6: Size and Power Using 5% Fixed-$b$ Critical Values with Traditional Lag Choices, $T = 50$ and 100

| **T=50** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_0$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_0$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_0 \times l_0$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{12}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.053 | 0.052 | 0.053 | 0.057 | 0.959 | 0.597 | 0.085 | 0.031 | 0.050 | 0.025 | 0.005 | 0.003 |
| 0.25 | 0.149 | 0.073 | 0.061 | 0.061 | 0.923 | 0.608 | 0.096 | 0.034 | 0.134 | 0.035 | 0.005 | 0.002 |
| 0.5 | 0.336 | 0.114 | 0.071 | 0.065 | 0.867 | 0.613 | 0.121 | 0.042 | 0.280 | 0.056 | 0.007 | 0.002 |
| 0.75 | 0.653 | 0.236 | 0.102 | 0.066 | 0.641 | 0.487 | 0.168 | 0.058 | 0.379 | 0.073 | 0.010 | 0.002 |
| 0.95 | 0.924 | 0.551 | 0.247 | 0.079 | 0.108 | 0.103 | 0.083 | 0.057 | 0.082 | 0.016 | 0.006 | 0.000 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 0.961 | 0.727 | 0.495 | 0.168 | 0.042 | 0.041 | 0.039 | 0.039 | 0.030 | 0.007 | 0.003 | 0.001 |
| 0.25 | 0.944 | 0.717 | 0.492 | 0.170 | 0.141 | 0.059 | 0.039 | 0.037 | 0.112 | 0.012 | 0.003 | 0.001 |
| 0.5 | 0.895 | 0.696 | 0.481 | 0.168 | 0.400 | 0.113 | 0.037 | 0.030 | 0.321 | 0.031 | 0.003 | 0.001 |
| 0.75 | 0.709 | 0.589 | 0.414 | 0.154 | 0.786 | 0.321 | 0.044 | 0.026 | 0.518 | 0.129 | 0.003 | 0.005 |
| 0.95 | 0.126 | 0.114 | 0.097 | 0.079 | 0.955 | 0.586 | 0.077 | 0.033 | 0.117 | 0.058 | 0.005 | 0.004 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d$ =0.1 | 0.133 | 0.095 | 0.072 | 0.065 | 0.946 | 0.578 | 0.085 | 0.034 | 0.124 | 0.047 | 0.005 | 0.003 |
| 0.2 | 0.252 | 0.156 | 0.097 | 0.073 | 0.925 | 0.556 | 0.082 | 0.035 | 0.227 | 0.075 | 0.005 | 0.003 |
| 0.3 | 0.399 | 0.222 | 0.128 | 0.083 | 0.890 | 0.508 | 0.081 | 0.037 | 0.340 | 0.089 | 0.005 | 0.002 |
| 0.4 | 0.546 | 0.301 | 0.162 | 0.089 | 0.820 | 0.443 | 0.083 | 0.033 | 0.427 | 0.093 | 0.005 | 0.002 |
| 0.45 | 0.616 | 0.348 | 0.185 | 0.094 | 0.774 | 0.409 | 0.081 | 0.031 | 0.450 | 0.093 | 0.005 | 0.002 |
| 0.499 | 0.674 | 0.397 | 0.208 | 0.101 | 0.723 | 0.373 | 0.078 | 0.032 | 0.458 | 0.092 | 0.006 | 0.002 |
| 0.5 | 0.679 | 0.396 | 0.210 | 0.101 | 0.720 | 0.359 | 0.075 | 0.030 | 0.457 | 0.091 | 0.006 | 0.002 |
| 0.6 | 0.775 | 0.484 | 0.257 | 0.113 | 0.572 | 0.265 | 0.072 | 0.034 | 0.402 | 0.069 | 0.005 | 0.002 |
| 0.7 | 0.851 | 0.556 | 0.316 | 0.126 | 0.389 | 0.181 | 0.064 | 0.036 | 0.290 | 0.044 | 0.005 | 0.002 |
| 0.75 | 0.879 | 0.593 | 0.343 | 0.132 | 0.294 | 0.145 | 0.061 | 0.036 | 0.222 | 0.033 | 0.005 | 0.002 |
| 0.8 | 0.907 | 0.623 | 0.374 | 0.137 | 0.215 | 0.119 | 0.056 | 0.036 | 0.164 | 0.027 | 0.004 | 0.001 |
| 0.9 | 0.939 | 0.681 | 0.436 | 0.155 | 0.102 | 0.073 | 0.045 | 0.039 | 0.077 | 0.014 | 0.003 | 0.001 |

| **T=100** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ | $l_{50} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.045 | 0.045 | 0.048 | 0.051 | 0.801 | 0.391 | 0.054 | 0.028 | 0.032 | 0.013 | 0.002 | 0.002 |
| 0.25 | 0.060 | 0.052 | 0.052 | 0.053 | 0.813 | 0.410 | 0.057 | 0.028 | 0.044 | 0.015 | 0.003 | 0.002 |
| 0.5 | 0.095 | 0.061 | 0.057 | 0.058 | 0.841 | 0.447 | 0.072 | 0.032 | 0.073 | 0.016 | 0.004 | 0.002 |
| 0.75 | 0.208 | 0.093 | 0.068 | 0.065 | 0.839 | 0.508 | 0.109 | 0.042 | 0.158 | 0.023 | 0.010 | 0.001 |
| 0.95 | 0.608 | 0.305 | 0.137 | 0.054 | 0.244 | 0.209 | 0.132 | 0.072 | 0.092 | 0.015 | 0.012 | 0.002 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 0.832 | 0.621 | 0.445 | 0.086 | 0.042 | 0.043 | 0.044 | 0.044 | 0.017 | 0.004 | 0.005 | 0.002 |
| 0.25 | 0.830 | 0.617 | 0.444 | 0.085 | 0.059 | 0.046 | 0.042 | 0.040 | 0.026 | 0.005 | 0.005 | 0.002 |
| 0.5 | 0.820 | 0.615 | 0.442 | 0.084 | 0.123 | 0.063 | 0.040 | 0.038 | 0.060 | 0.008 | 0.004 | 0.002 |
| 0.75 | 0.758 | 0.588 | 0.422 | 0.081 | 0.389 | 0.123 | 0.038 | 0.036 | 0.238 | 0.027 | 0.003 | 0.004 |
| 0.95 | 0.277 | 0.236 | 0.182 | 0.079 | 0.777 | 0.372 | 0.051 | 0.030 | 0.203 | 0.072 | 0.003 | 0.007 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d$ =0.1 | 0.100 | 0.077 | 0.065 | 0.057 | 0.788 | 0.393 | 0.054 | 0.028 | 0.073 | 0.023 | 0.020 | 0.003 |
| 0.2 | 0.186 | 0.121 | 0.085 | 0.062 | 0.767 | 0.379 | 0.055 | 0.029 | 0.134 | 0.030 | 0.025 | 0.003 |
| 0.3 | 0.288 | 0.171 | 0.110 | 0.062 | 0.726 | 0.362 | 0.059 | 0.030 | 0.197 | 0.035 | 0.026 | 0.003 |
| 0.4 | 0.390 | 0.231 | 0.146 | 0.074 | 0.667 | 0.314 | 0.060 | 0.032 | 0.235 | 0.032 | 0.023 | 0.001 |
| 0.45 | 0.439 | 0.259 | 0.163 | 0.077 | 0.623 | 0.285 | 0.060 | 0.033 | 0.243 | 0.035 | 0.023 | 0.002 |
| 0.499 | 0.489 | 0.292 | 0.182 | 0.078 | 0.576 | 0.246 | 0.062 | 0.036 | 0.249 | 0.030 | 0.017 | 0.002 |
| 0.5 | 0.489 | 0.292 | 0.184 | 0.076 | 0.577 | 0.254 | 0.057 | 0.034 | 0.246 | 0.034 | 0.018 | 0.002 |
| 0.6 | 0.575 | 0.360 | 0.220 | 0.080 | 0.434 | 0.193 | 0.060 | 0.038 | 0.196 | 0.022 | 0.014 | 0.002 |
| 0.7 | 0.655 | 0.428 | 0.266 | 0.078 | 0.294 | 0.143 | 0.060 | 0.038 | 0.131 | 0.015 | 0.010 | 0.003 |
| 0.75 | 0.691 | 0.459 | 0.290 | 0.080 | 0.230 | 0.120 | 0.057 | 0.039 | 0.102 | 0.010 | 0.008 | 0.002 |
| 0.8 | 0.728 | 0.494 | 0.316 | 0.081 | 0.172 | 0.098 | 0.057 | 0.039 | 0.076 | 0.008 | 0.006 | 0.002 |
| 0.9 | 0.786 | 0.562 | 0.383 | 0.083 | 0.091 | 0.065 | 0.050 | 0.041 | 0.037 | 0.006 | 0.004 | 0.002 |

Table 3.7: Size and Power Using 5% Fixed-$b$ Critical Values with Traditional Lag Choices, $T =$ 200 and 500

| **T=200** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ | $l_{50} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.047 | 0.053 | 0.053 | 0.053 | 0.956 | 0.663 | 0.346 | 0.046 | 0.045 | 0.034 | 0.017 | 0.017 |
| 0.25 | 0.066 | 0.057 | 0.055 | 0.056 | 0.963 | 0.677 | 0.356 | 0.049 | 0.063 | 0.038 | 0.016 | 0.018 |
| 0.5 | 0.102 | 0.065 | 0.060 | 0.057 | 0.974 | 0.727 | 0.395 | 0.060 | 0.100 | 0.044 | 0.019 | 0.021 |
| 0.75 | 0.233 | 0.094 | 0.074 | 0.064 | 0.983 | 0.826 | 0.475 | 0.084 | 0.227 | 0.074 | 0.023 | 0.025 |
| 0.95 | 0.720 | 0.340 | 0.186 | 0.091 | 0.641 | 0.540 | 0.407 | 0.150 | 0.414 | 0.119 | 0.020 | 0.017 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 0.951 | 0.741 | 0.587 | 0.377 | 0.048 | 0.047 | 0.048 | 0.046 | 0.034 | 0.011 | 0.002 | 0.002 |
| 0.25 | 0.950 | 0.739 | 0.587 | 0.378 | 0.073 | 0.051 | 0.051 | 0.046 | 0.054 | 0.013 | 0.003 | 0.002 |
| 0.50 | 0.948 | 0.738 | 0.583 | 0.378 | 0.150 | 0.069 | 0.053 | 0.046 | 0.122 | 0.022 | 0.004 | 0.003 |
| 0.75 | 0.923 | 0.725 | 0.576 | 0.376 | 0.478 | 0.180 | 0.083 | 0.041 | 0.419 | 0.078 | 0.010 | 0.006 |
| 0.95 | 0.564 | 0.487 | 0.414 | 0.283 | 0.930 | 0.616 | 0.286 | 0.046 | 0.516 | 0.287 | 0.095 | 0.060 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d = 0.1$ | 0.137 | 0.099 | 0.084 | 0.072 | 0.952 | 0.661 | 0.344 | 0.048 | 0.128 | 0.065 | 0.054 | 0.027 |
| 0.2 | 0.271 | 0.167 | 0.123 | 0.091 | 0.941 | 0.652 | 0.334 | 0.049 | 0.252 | 0.109 | 0.079 | 0.035 |
| 0.3 | 0.415 | 0.241 | 0.173 | 0.115 | 0.920 | 0.624 | 0.318 | 0.051 | 0.375 | 0.143 | 0.102 | 0.038 |
| 0.4 | 0.550 | 0.318 | 0.221 | 0.141 | 0.884 | 0.575 | 0.285 | 0.053 | 0.476 | 0.161 | 0.107 | 0.034 |
| 0.45 | 0.608 | 0.364 | 0.246 | 0.152 | 0.852 | 0.531 | 0.262 | 0.055 | 0.504 | 0.165 | 0.104 | 0.028 |
| 0.499 | 0.655 | 0.406 | 0.281 | 0.174 | 0.811 | 0.492 | 0.239 | 0.057 | 0.513 | 0.158 | 0.102 | 0.024 |
| 0.5 | 0.659 | 0.408 | 0.281 | 0.175 | 0.811 | 0.492 | 0.237 | 0.053 | 0.516 | 0.163 | 0.104 | 0.030 |
| 0.6 | 0.749 | 0.481 | 0.343 | 0.204 | 0.677 | 0.375 | 0.188 | 0.058 | 0.478 | 0.126 | 0.078 | 0.018 |
| 0.7 | 0.826 | 0.550 | 0.398 | 0.238 | 0.479 | 0.258 | 0.142 | 0.060 | 0.357 | 0.083 | 0.047 | 0.011 |
| 0.75 | 0.860 | 0.584 | 0.428 | 0.257 | 0.377 | 0.201 | 0.122 | 0.059 | 0.289 | 0.060 | 0.034 | 0.009 |
| 0.8 | 0.884 | 0.617 | 0.459 | 0.278 | 0.280 | 0.158 | 0.101 | 0.055 | 0.216 | 0.045 | 0.025 | 0.007 |
| 0.9 | 0.925 | 0.684 | 0.518 | 0.323 | 0.126 | 0.086 | 0.071 | 0.053 | 0.094 | 0.021 | 0.013 | 0.004 |

| **T=500** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ | $l_{50} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.054 | 0.055 | 0.054 | 0.052 | 0.998 | 0.901 | 0.672 | 0.365 | 0.054 | 0.049 | 0.035 | 0.033 |
| 0.25 | 0.070 | 0.058 | 0.056 | 0.054 | 0.998 | 0.908 | 0.687 | 0.377 | 0.069 | 0.052 | 0.037 | 0.035 |
| 0.5 | 0.097 | 0.064 | 0.060 | 0.056 | 1.000 | 0.941 | 0.738 | 0.415 | 0.097 | 0.059 | 0.042 | 0.039 |
| 0.75 | 0.197 | 0.090 | 0.068 | 0.063 | 1.000 | 0.983 | 0.848 | 0.507 | 0.197 | 0.088 | 0.055 | 0.049 |
| 0.95 | 0.733 | 0.335 | 0.166 | 0.099 | 0.997 | 0.984 | 0.927 | 0.639 | 0.730 | 0.325 | 0.139 | 0.084 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 0.992 | 0.908 | 0.749 | 0.593 | 0.048 | 0.051 | 0.047 | 0.049 | 0.044 | 0.028 | 0.013 | 0.007 |
| 0.25 | 0.992 | 0.908 | 0.748 | 0.593 | 0.067 | 0.055 | 0.049 | 0.048 | 0.062 | 0.031 | 0.014 | 0.008 |
| 0.5 | 0.992 | 0.907 | 0.748 | 0.592 | 0.131 | 0.070 | 0.057 | 0.050 | 0.126 | 0.043 | 0.017 | 0.009 |
| 0.75 | 0.990 | 0.906 | 0.747 | 0.593 | 0.454 | 0.193 | 0.101 | 0.060 | 0.445 | 0.146 | 0.037 | 0.021 |
| 0.95 | 0.890 | 0.802 | 0.685 | 0.552 | 0.984 | 0.815 | 0.545 | 0.227 | 0.874 | 0.641 | 0.347 | 0.269 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d = 0.1$ | 0.178 | 0.129 | 0.102 | 0.085 | 0.997 | 0.897 | 0.669 | 0.363 | 0.178 | 0.114 | 0.090 | 0.066 |
| 0.2 | 0.371 | 0.239 | 0.164 | 0.123 | 0.996 | 0.888 | 0.662 | 0.354 | 0.369 | 0.210 | 0.144 | 0.104 |
| 0.3 | 0.559 | 0.356 | 0.244 | 0.172 | 0.992 | 0.871 | 0.647 | 0.336 | 0.554 | 0.304 | 0.207 | 0.147 |
| 0.4 | 0.725 | 0.478 | 0.324 | 0.220 | 0.983 | 0.835 | 0.598 | 0.290 | 0.712 | 0.389 | 0.264 | 0.178 |
| 0.45 | 0.790 | 0.530 | 0.365 | 0.244 | 0.970 | 0.803 | 0.564 | 0.274 | 0.762 | 0.415 | 0.280 | 0.182 |
| 0.499 | 0.836 | 0.583 | 0.408 | 0.278 | 0.954 | 0.755 | 0.524 | 0.242 | 0.794 | 0.422 | 0.288 | 0.186 |
| 0.5 | 0.834 | 0.582 | 0.409 | 0.278 | 0.952 | 0.757 | 0.516 | 0.245 | 0.790 | 0.421 | 0.287 | 0.176 |
| 0.6 | 0.909 | 0.677 | 0.489 | 0.339 | 0.855 | 0.619 | 0.390 | 0.181 | 0.770 | 0.386 | 0.264 | 0.138 |
| 0.7 | 0.952 | 0.758 | 0.559 | 0.405 | 0.658 | 0.419 | 0.253 | 0.134 | 0.617 | 0.269 | 0.178 | 0.085 |
| 0.75 | 0.965 | 0.790 | 0.596 | 0.436 | 0.523 | 0.322 | 0.196 | 0.112 | 0.493 | 0.207 | 0.135 | 0.065 |
| 0.8 | 0.975 | 0.819 | 0.636 | 0.464 | 0.387 | 0.233 | 0.146 | 0.096 | 0.367 | 0.148 | 0.094 | 0.046 |
| 0.9 | 0.988 | 0.871 | 0.693 | 0.530 | 0.153 | 0.107 | 0.084 | 0.069 | 0.146 | 0.064 | 0.036 | 0.022 |

Table 3.8: Size and Power Using 5% Fixed-$b$ Critical Values with Traditional Lag Choices, $T =$ 1,000 and 2,000

| **T=1,000** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ | $l_{50} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.049 | 0.049 | 0.051 | 0.049 | 1.000 | 0.973 | 0.851 | 0.621 | 0.049 | 0.048 | 0.044 | 0.042 |
| 0.25 | 0.059 | 0.054 | 0.052 | 0.050 | 1.000 | 0.978 | 0.866 | 0.635 | 0.059 | 0.052 | 0.045 | 0.043 |
| 0.5 | 0.079 | 0.059 | 0.056 | 0.052 | 1.000 | 0.989 | 0.901 | 0.685 | 0.079 | 0.059 | 0.051 | 0.046 |
| 0.75 | 0.151 | 0.077 | 0.063 | 0.055 | 1.000 | 0.999 | 0.966 | 0.807 | 0.151 | 0.077 | 0.061 | 0.053 |
| 0.95 | 0.647 | 0.286 | 0.145 | 0.092 | 1.000 | 1.000 | 0.998 | 0.964 | 0.647 | 0.286 | 0.144 | 0.091 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 0.998 | 0.963 | 0.861 | 0.707 | 0.052 | 0.053 | 0.052 | 0.054 | 0.051 | 0.040 | 0.023 | 0.012 |
| 0.25 | 0.998 | 0.962 | 0.860 | 0.707 | 0.065 | 0.057 | 0.053 | 0.055 | 0.063 | 0.044 | 0.023 | 0.012 |
| 0.5 | 0.998 | 0.962 | 0.860 | 0.707 | 0.123 | 0.076 | 0.064 | 0.058 | 0.121 | 0.059 | 0.029 | 0.015 |
| 0.75 | 0.997 | 0.961 | 0.860 | 0.707 | 0.396 | 0.195 | 0.117 | 0.079 | 0.393 | 0.170 | 0.066 | 0.036 |
| 0.95 | 0.976 | 0.934 | 0.832 | 0.686 | 0.988 | 0.867 | 0.659 | 0.413 | 0.965 | 0.805 | 0.528 | 0.419 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d$ =0.1 | 0.201 | 0.146 | 0.116 | 0.094 | 1.000 | 0.972 | 0.848 | 0.620 | 0.201 | 0.140 | 0.099 | 0.080 |
| 0.2 | 0.418 | 0.281 | 0.207 | 0.150 | 1.000 | 0.968 | 0.844 | 0.616 | 0.418 | 0.271 | 0.172 | 0.126 |
| 0.3 | 0.635 | 0.432 | 0.316 | 0.221 | 0.999 | 0.960 | 0.829 | 0.601 | 0.635 | 0.413 | 0.257 | 0.180 |
| 0.4 | 0.803 | 0.584 | 0.425 | 0.296 | 0.998 | 0.935 | 0.800 | 0.570 | 0.802 | 0.541 | 0.330 | 0.227 |
| 0.45 | 0.862 | 0.647 | 0.476 | 0.334 | 0.994 | 0.916 | 0.764 | 0.531 | 0.857 | 0.584 | 0.350 | 0.240 |
| 0.499 | 0.901 | 0.703 | 0.527 | 0.378 | 0.987 | 0.886 | 0.716 | 0.486 | 0.889 | 0.613 | 0.357 | 0.253 |
| 0.5 | 0.902 | 0.707 | 0.528 | 0.378 | 0.986 | 0.885 | 0.711 | 0.485 | 0.888 | 0.614 | 0.351 | 0.246 |
| 0.6 | 0.955 | 0.799 | 0.619 | 0.456 | 0.929 | 0.754 | 0.573 | 0.366 | 0.884 | 0.581 | 0.313 | 0.217 |
| 0.7 | 0.978 | 0.862 | 0.693 | 0.529 | 0.749 | 0.547 | 0.392 | 0.248 | 0.728 | 0.443 | 0.213 | 0.146 |
| 0.75 | 0.987 | 0.890 | 0.728 | 0.563 | 0.615 | 0.431 | 0.307 | 0.200 | 0.603 | 0.356 | 0.163 | 0.108 |
| 0.8 | 0.991 | 0.911 | 0.761 | 0.592 | 0.455 | 0.319 | 0.227 | 0.153 | 0.448 | 0.263 | 0.121 | 0.075 |
| 0.9 | 0.996 | 0.943 | 0.815 | 0.647 | 0.194 | 0.143 | 0.115 | 0.093 | 0.190 | 0.115 | 0.055 | 0.030 |

| **T=2,000** | $\widehat{\eta}_\mu$ | | | | $\widehat{\eta}_\mu^d$ | | | | $\widehat{\eta}_\mu \times \widehat{\eta}_\mu^d$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lag* | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4$ | $l_{12}$ | $l_{25}$ | $l_{50}$ | $l_4 \times l_4$ | $l_{12} \times l_{12}$ | $l_{25} \times l_{25}$ | $l_{50} \times l_{25}$ |
| $\epsilon_t \sim I(0)$ | (Size) | | | | (Power) | | | | (Size) | | | |
| $\rho = 0$ | 0.049 | 0.049 | 0.047 | 0.047 | 1.000 | 0.996 | 0.954 | 0.802 | 0.049 | 0.049 | 0.045 | 0.045 |
| 0.25 | 0.060 | 0.051 | 0.049 | 0.048 | 1.000 | 0.997 | 0.960 | 0.818 | 0.060 | 0.051 | 0.047 | 0.046 |
| 0.5 | 0.074 | 0.056 | 0.052 | 0.049 | 1.000 | 0.999 | 0.979 | 0.866 | 0.074 | 0.056 | 0.051 | 0.049 |
| 0.75 | 0.134 | 0.070 | 0.057 | 0.052 | 1.000 | 1.000 | 0.996 | 0.950 | 0.134 | 0.070 | 0.057 | 0.052 |
| 0.95 | 0.604 | 0.245 | 0.133 | 0.080 | 1.000 | 1.000 | 1.000 | 1.000 | 0.604 | 0.245 | 0.133 | 0.080 |
| $\epsilon_t \sim I(1)$ | (Power) | | | | (Size) | | | | (Size) | | | |
| $\phi = 0$ | 1.000 | 0.990 | 0.944 | 0.815 | 0.052 | 0.051 | 0.052 | 0.053 | 0.051 | 0.048 | 0.037 | 0.019 |
| 0.25 | 1.000 | 0.990 | 0.944 | 0.815 | 0.065 | 0.058 | 0.054 | 0.054 | 0.064 | 0.054 | 0.038 | 0.020 |
| 0.5 | 1.000 | 0.990 | 0.943 | 0.815 | 0.114 | 0.073 | 0.061 | 0.058 | 0.114 | 0.069 | 0.044 | 0.024 |
| 0.75 | 1.000 | 0.989 | 0.943 | 0.815 | 0.380 | 0.176 | 0.108 | 0.078 | 0.380 | 0.171 | 0.084 | 0.051 |
| 0.95 | 0.998 | 0.983 | 0.932 | 0.810 | 0.995 | 0.886 | 0.708 | 0.482 | 0.993 | 0.870 | 0.648 | 0.546 |
| $\epsilon_t \sim I(d)$ | (Power) | | | | (Power) | | | | (Power) | | | |
| $d$ =0.1 | 0.226 | 0.167 | 0.131 | 0.107 | 1.000 | 0.996 | 0.953 | 0.800 | 0.226 | 0.166 | 0.124 | 0.103 |
| 0.2 | 0.499 | 0.350 | 0.267 | 0.191 | 1.000 | 0.995 | 0.950 | 0.799 | 0.499 | 0.349 | 0.254 | 0.182 |
| 0.3 | 0.749 | 0.538 | 0.407 | 0.287 | 1.000 | 0.993 | 0.940 | 0.786 | 0.749 | 0.534 | 0.382 | 0.271 |
| 0.4 | 0.892 | 0.700 | 0.535 | 0.393 | 1.000 | 0.988 | 0.912 | 0.755 | 0.892 | 0.691 | 0.484 | 0.354 |
| 0.45 | 0.935 | 0.770 | 0.599 | 0.434 | 1.000 | 0.979 | 0.892 | 0.723 | 0.934 | 0.754 | 0.529 | 0.381 |
| 0.499 | 0.959 | 0.819 | 0.662 | 0.487 | 0.999 | 0.961 | 0.856 | 0.677 | 0.958 | 0.785 | 0.554 | 0.405 |
| 0.5 | 0.960 | 0.824 | 0.663 | 0.489 | 0.999 | 0.961 | 0.852 | 0.675 | 0.959 | 0.790 | 0.551 | 0.403 |
| 0.6 | 0.988 | 0.893 | 0.759 | 0.570 | 0.978 | 0.869 | 0.723 | 0.538 | 0.966 | 0.767 | 0.523 | 0.380 |
| 0.7 | 0.995 | 0.942 | 0.826 | 0.651 | 0.853 | 0.669 | 0.526 | 0.369 | 0.848 | 0.617 | 0.397 | 0.293 |
| 0.75 | 0.996 | 0.954 | 0.855 | 0.688 | 0.712 | 0.543 | 0.408 | 0.279 | 0.709 | 0.504 | 0.308 | 0.224 |
| 0.8 | 0.998 | 0.966 | 0.878 | 0.718 | 0.557 | 0.408 | 0.295 | 0.211 | 0.556 | 0.381 | 0.222 | 0.155 |
| 0.9 | 1.000 | 0.982 | 0.916 | 0.769 | 0.222 | 0.167 | 0.136 | 0.110 | 0.221 | 0.158 | 0.099 | 0.061 |

Table 3.9: Size and Power of $ADF$ and $\widehat{\eta}_{\mu} \times ADF$ Tests Using Standard Critical Values with Traditional Lag Choices, $T = 50, 100, 200,$ and 500

| | T=50 | | | | | | T=100 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ADF$ | | | $\widehat{\eta}_{\mu} \times ADF$ | | | $ADF$ | | | $\widehat{\eta}_{\mu} \times ADF$ | | |
| $lag$ | $p_0$ | $p_4$ | $p_{12}$ | $l_0 \times p_0$ | $l_4 \times p_4$ | $l_{12} \times p_{12}$ | $p_0$ | $p_4$ | $p_{12}$ | $l_0 \times p_0$ | $l_4 \times p_4$ | $l_{12} \times p_{12}$ |
| $\epsilon_t \sim I(0)$ | | (Power) | | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 1.000 | 0.789 | 0.072 | 0.053 | 0.000 | 0.000 | 1.000 | 0.994 | 0.346 | 0.043 | 0.002 | 0.007 |
| 0.25 | 1.000 | 0.683 | 0.069 | 0.149 | 0.000 | 0.000 | 1.000 | 0.983 | 0.315 | 0.143 | 0.001 | 0.008 |
| 0.50 | 0.978 | 0.495 | 0.061 | 0.320 | 0.001 | 0.000 | 1.000 | 0.932 | 0.270 | 0.361 | 0.003 | 0.010 |
| 0.75 | 0.474 | 0.220 | 0.047 | 0.238 | 0.002 | 0.000 | 0.974 | 0.623 | 0.185 | 0.693 | 0.008 | 0.011 |
| 0.95 | 0.064 | 0.056 | 0.038 | 0.043 | 0.012 | 0.000 | 0.116 | 0.091 | 0.060 | 0.105 | 0.013 | 0.002 |
| $\epsilon_t \sim I(1)$ | | (Size) | | | (Size) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.041 | 0.044 | 0.038 | 0.032 | 0.020 | 0.000 | 0.054 | 0.051 | 0.043 | 0.051 | 0.023 | 0.004 |
| 0.25 | 0.157 | 0.041 | 0.034 | 0.119 | 0.018 | 0.000 | 0.190 | 0.050 | 0.044 | 0.179 | 0.024 | 0.011 |
| 0.50 | 0.515 | 0.052 | 0.031 | 0.413 | 0.014 | 0.000 | 0.585 | 0.058 | 0.041 | 0.563 | 0.022 | 0.056 |
| 0.75 | 0.956 | 0.182 | 0.030 | 0.665 | 0.010 | 0.000 | 0.978 | 0.182 | 0.041 | 0.883 | 0.016 | 0.170 |
| 0.95 | 1.000 | 0.716 | 0.065 | 0.126 | 0.001 | 0.000 | 1.000 | 0.947 | 0.206 | 0.300 | 0.008 | 0.049 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 1.000 | 0.672 | 0.063 | 0.133 | 0.001 | 0.000 | 1.000 | 0.970 | 0.274 | 0.156 | 0.002 | 0.011 |
| 0.2 | 1.000 | 0.536 | 0.057 | 0.252 | 0.001 | 0.000 | 1.000 | 0.901 | 0.205 | 0.346 | 0.005 | 0.017 |
| 0.3 | 0.999 | 0.406 | 0.050 | 0.398 | 0.002 | 0.000 | 1.000 | 0.761 | 0.158 | 0.539 | 0.009 | 0.025 |
| 0.4 | 0.978 | 0.294 | 0.043 | 0.525 | 0.004 | 0.000 | 1.000 | 0.573 | 0.123 | 0.715 | 0.012 | 0.034 |
| 0.45 | 0.941 | 0.249 | 0.043 | 0.559 | 0.004 | 0.000 | 1.000 | 0.479 | 0.111 | 0.782 | 0.013 | 0.040 |
| 0.499 | 0.882 | 0.209 | 0.042 | 0.561 | 0.007 | 0.000 | 0.998 | 0.401 | 0.101 | 0.828 | 0.017 | 0.048 |
| 0.5 | 0.880 | 0.207 | 0.042 | 0.564 | 0.007 | 0.000 | 0.998 | 0.397 | 0.100 | 0.828 | 0.017 | 0.048 |
| 0.6 | 0.686 | 0.143 | 0.040 | 0.476 | 0.010 | 0.000 | 0.930 | 0.255 | 0.082 | 0.838 | 0.019 | 0.053 |
| 0.7 | 0.422 | 0.102 | 0.037 | 0.298 | 0.013 | 0.000 | 0.692 | 0.155 | 0.067 | 0.645 | 0.019 | 0.038 |
| 0.75 | 0.307 | 0.084 | 0.039 | 0.217 | 0.015 | 0.000 | 0.516 | 0.128 | 0.059 | 0.484 | 0.020 | 0.028 |
| 0.8 | 0.219 | 0.070 | 0.038 | 0.159 | 0.017 | 0.000 | 0.362 | 0.102 | 0.055 | 0.338 | 0.021 | 0.019 |
| 0.9 | 0.094 | 0.053 | 0.037 | 0.068 | 0.017 | 0.000 | 0.145 | 0.068 | 0.048 | 0.133 | 0.024 | 0.006 |

| | T=200 | | | | | | T=500 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ADF$ | | | $\widehat{\eta}_{\mu} \times ADF$ | | | $ADF$ | | | $\widehat{\eta}_{\mu} \times ADF$ | | |
| $lag$ | $p_4$ | $p_{12}$ | $p_{25}$ | $l_4 \times p_4$ | $l_{12} \times p_{12}$ | $l_{25} \times p_{25}$ | $p_4$ | $p_{12}$ | $p_{25}$ | $l_4 \times p_4$ | $l_{12} \times p_{12}$ | $l_{25} \times p_{25}$ |
| $\epsilon_t \sim I(0)$ | | (Power) | | | (Size) | | | (Power) | | | (Size) | |
| $\rho = 0$ | 1.000 | 1.000 | 0.875 | 0.054 | 0.043 | 0.028 | 1.000 | 1.000 | 0.875 | 0.054 | 0.043 | 0.028 |
| 0.25 | 1.000 | 1.000 | 0.867 | 0.067 | 0.044 | 0.028 | 1.000 | 1.000 | 0.867 | 0.067 | 0.044 | 0.028 |
| 0.50 | 1.000 | 1.000 | 0.846 | 0.094 | 0.047 | 0.030 | 1.000 | 1.000 | 0.846 | 0.094 | 0.047 | 0.030 |
| 0.75 | 1.000 | 0.999 | 0.778 | 0.192 | 0.056 | 0.033 | 1.000 | 0.999 | 0.778 | 0.192 | 0.056 | 0.033 |
| 0.95 | 0.899 | 0.661 | 0.359 | 0.422 | 0.135 | 0.030 | 0.899 | 0.661 | 0.359 | 0.422 | 0.135 | 0.030 |
| $\epsilon_t \sim I(1)$ | | (Size) | | | (Size) | | | (Size) | | | (Size) | |
| $\phi = 0$ | 0.046 | 0.047 | 0.046 | 0.043 | 0.018 | 0.011 | 0.046 | 0.047 | 0.046 | 0.043 | 0.018 | 0.011 |
| 0.25 | 0.046 | 0.045 | 0.047 | 0.042 | 0.073 | 0.011 | 0.046 | 0.045 | 0.047 | 0.042 | 0.073 | 0.011 |
| 0.50 | 0.052 | 0.045 | 0.048 | 0.041 | 0.359 | 0.010 | 0.052 | 0.045 | 0.048 | 0.041 | 0.359 | 0.010 |
| 0.75 | 0.165 | 0.047 | 0.047 | 0.043 | 0.719 | 0.011 | 0.165 | 0.047 | 0.047 | 0.043 | 0.719 | 0.011 |
| 0.95 | 0.997 | 0.480 | 0.098 | 0.371 | 0.661 | 0.121 | 0.997 | 0.480 | 0.098 | 0.371 | 0.661 | 0.121 |
| $\epsilon_t \sim I(d)$ | | (Power) | | | (Power) | | | (Power) | | | (Power) | |
| $d$ =0.1 | 1.000 | 0.751 | 0.193 | 0.045 | 0.051 | 0.004 | 1.000 | 1.000 | 0.749 | 0.173 | 0.084 | 0.050 |
| 0.2 | 1.000 | 0.595 | 0.149 | 0.066 | 0.082 | 0.005 | 1.000 | 0.993 | 0.592 | 0.355 | 0.145 | 0.077 |
| 0.3 | 0.997 | 0.444 | 0.116 | 0.077 | 0.117 | 0.005 | 1.000 | 0.944 | 0.432 | 0.484 | 0.220 | 0.094 |
| 0.4 | 0.956 | 0.313 | 0.089 | 0.064 | 0.161 | 0.004 | 0.999 | 0.798 | 0.300 | 0.498 | 0.297 | 0.091 |
| 0.45 | 0.892 | 0.265 | 0.081 | 0.064 | 0.183 | 0.003 | 0.999 | 0.695 | 0.252 | 0.459 | 0.339 | 0.083 |
| 0.499 | 0.791 | 0.225 | 0.077 | 0.063 | 0.214 | 0.003 | 0.991 | 0.575 | 0.206 | 0.384 | 0.379 | 0.072 |
| 0.5 | 0.788 | 0.224 | 0.076 | 0.062 | 0.215 | 0.003 | 0.990 | 0.572 | 0.205 | 0.378 | 0.381 | 0.069 |
| 0.6 | 0.543 | 0.153 | 0.060 | 0.051 | 0.267 | 0.002 | 0.870 | 0.356 | 0.139 | 0.253 | 0.461 | 0.040 |
| 0.7 | 0.317 | 0.110 | 0.052 | 0.044 | 0.249 | 0.002 | 0.552 | 0.203 | 0.094 | 0.154 | 0.516 | 0.022 |
| 0.75 | 0.231 | 0.094 | 0.049 | 0.043 | 0.184 | 0.002 | 0.398 | 0.152 | 0.083 | 0.115 | 0.492 | 0.018 |
| 0.8 | 0.169 | 0.079 | 0.046 | 0.040 | 0.110 | 0.001 | 0.262 | 0.115 | 0.068 | 0.094 | 0.367 | 0.013 |
| 0.9 | 0.096 | 0.060 | 0.044 | 0.039 | 0.034 | 0.002 | 0.106 | 0.071 | 0.055 | 0.062 | 0.095 | 0.010 |

APPENDIX

# Appendix for Chapter 3

**Proof of Theorem 8**

Liu (1998), Theorem 3.4, shows that $\sum S_t^2 = O_p\left(T^3\right)$ when $\epsilon_t$ is $I\left(\frac{1}{2}\right)$. It also shows that $s^2(0) = O_p(\ln T)$. However, he does not establish the order in probability of $s^2(l)$ when $l \to \infty$ as $T \to \infty$. The proof here will establish the limiting behavior of $s^2(l)$ when $l \to \infty$ using results of Tanaka (1999). Tanaka provides the invariance principle for $I\left(\frac{1}{2}\right)$ processes having i.i.d. innovations. He defines the process

$$X_T(t) = \frac{1}{s_T} y_j + \frac{ts_T^2 - s_j^2}{s_j^2 - s_{j-1}^2} \frac{1}{s_T}\left(y_j - y_{j-1}\right), \quad \left(\frac{s_{j-1}^2}{s_T^2} \leq t \leq \frac{s_j^2}{s_T^2}\right), \tag{2.37}$$

where $(1-L)^{1/2} y_t = u_t \sim IID(0, \sigma_u^2)$ and $s_j^2 = Var\left(y_j\right)$. Lemma 2.1 of Liu (1998) shows that

$$s_j^2 = \frac{4\sigma_u^2}{\pi} \sum_{k=1}^{j} \frac{1}{2k-1} = K \cdot L(j) \tag{2.38}$$

with $K = \frac{2\sigma_u^2}{\pi}$ and

$$\lim_{T\to\infty} \frac{L(T)}{\log T} = 1. \tag{2.39}$$

Theorem 2.1 in Tanaka (1999) states $X_T = \{X_T(t)\}$ weakly converges to the standard Wiener process defined on $[0,1]$. Note that in our terminology, $\epsilon_t$ replaces $y_t$ in Tanaka (1999). Rewrite $s^2(l)$ as

$$s^2(l) = \frac{1}{T}\sum_{t=1}^{T} e_t^2 + 2\sum_{s=1}^{l}\left(1 - \frac{s}{l+1}\right)\frac{1}{T}\sum_{t=s+1}^{T} e_t e_{t-s}.$$

Multiplying $s^2(l)$ by $\frac{1}{l \ln T} \cdot \frac{1}{\ln T}$ yields

$$\frac{1}{l \ln T} \cdot \frac{1}{\ln T} \cdot s^2(l) = \frac{1}{l} \frac{1}{\ln T} \frac{1}{T} \sum_{t=1}^{T} (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_t \qquad (2.40)$$
$$+ \frac{1}{l} \sum_{s=1}^{l} \left(1 - \frac{s}{l+1}\right) \frac{1}{\ln T} \frac{1}{T} \sum_{t=s+1}^{T} (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s}.$$

Consider the absolute value of the second sum.

$$\left| \frac{1}{l} \sum_{s=1}^{l} \left(1 - \frac{s}{l+1}\right) \frac{1}{\ln T} \frac{1}{T} \sum_{t=s+1}^{T} (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right| \qquad (2.41)$$

$$\leq \left( \frac{1}{l} \sum_{s=1}^{l} \left(1 - \frac{s}{l+1}\right) \right) \frac{1}{\ln T} \frac{1}{T} \cdot \max_{1 \leq s \leq l} \left| \sum_{t=s+1}^{T} (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right|$$
$$= \frac{1}{2} \frac{1}{\ln T} \frac{1}{T} \cdot \max_{1 \leq s \leq l} \left| \sum_{t=s+1}^{T} (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right|$$
$$\leq \frac{1}{2} \frac{1}{\ln T} \frac{1}{T} \cdot \max_{1 \leq s \leq l} \sum_{t=s+1}^{T} \left| (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right|.$$

It turns out that it is enough to show this last expression is $o_p(1)$.

$$\max_{1 \leq s \leq l} \left( \frac{1}{\ln T} \frac{1}{T} \sum_{t=s+1}^{T} \left| (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right| \right) \qquad (2.42)$$
$$\leq \max_{1 \leq s \leq l} \left( \frac{1}{\ln T} \frac{1}{T} (T-s) \max_{s+1 \leq t \leq T} \left| (\ln T)^{-1/2} e_t \cdot (\ln T)^{-1/2} e_{t-s} \right| \right)$$
$$\leq \max_{1 \leq s \leq l} \left( \frac{1}{\ln T} \max_{1 \leq t \leq T} \left| (\ln T)^{-1/2} e_t \right| \cdot \max_{1 \leq t \leq T} \left| (\ln T)^{-1/2} e_t \right| \right)$$
$$= \frac{1}{\ln T} \max_{1 \leq t \leq T} \left| (\ln T)^{-1/2} e_t \right| \cdot \max_{1 \leq t \leq T} \left| (\ln T)^{-1/2} e_t \right| = \frac{1}{\ln T} \cdot O_p(1) = o_p(1).$$

The second to last equality in (2.42) comes from the following.

$$\begin{aligned}\max_{1\le t\le T}\left|(\ln T)^{-1/2}e_t\right| &\le \max_{1\le t\le T}\left[\left|(\ln T)^{-1/2}\epsilon_t\right|+\left|(\ln T)^{-1/2}\overline{\epsilon}\right|\right] \quad (2.43)\\ &\le 2\cdot\max_{1\le t\le T}\left|(\ln T)^{-1/2}\epsilon_t\right|\simeq 2\cdot\max_{1\le t\le T}\left|\frac{K^{1/2}}{s_T}\cdot\epsilon_t\right| \text{ with large } T\\ &= 2K^{1/2}\cdot\max_{0\le r\le 1}|X_T(r)|\Rightarrow 2K^{1/2}\cdot\max_{0\le r\le 1}|W(r)|=O_p(1).\end{aligned}$$

The weak convergence result in the last line follows from Tanaka (1999).

Similarly, one can show $\frac{1}{l}\frac{1}{\ln T}\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sqrt{\ln T}}e_t\cdot\frac{1}{\sqrt{\ln T}}e_t = o_p(1)$. Hence $\frac{1}{l(\ln T)^2}\cdot s^2(l) = o_p(1)$. Now rewrite $\widehat{\eta}_\mu$ as

$$\widehat{\eta}_\mu = \frac{\frac{1}{T^3}\sum_{t=1}^{T}S_t^2}{\frac{1}{l(\ln T)^2}s^2(l)}\times\frac{T}{l\,(\ln T)^2}, \quad (2.44)$$

and recall that $\frac{1}{T^3}\sum_{t=1}^{T}S_t^2 = O_p(1)$ and its weak limit is not zero (Liu 1998). Also from the above $p\lim\frac{1}{l(\ln T)^2}s^2(l) = 0$. Therefore, since $\frac{T}{l(\ln T)^2}$ goes to infinity under the traditional choice of the number of the lags, $\widehat{\eta}_\mu$ diverges to infinity if $l\to\infty$ and $\frac{l}{T}\to 0$ as $T\to\infty$.

**Proof of Proposition 8**

In this case, $\Delta\epsilon_t$ is a short memory process with zero mean. So the limiting behavior of $\widetilde{s}^2(l)$ and $\widetilde{S}_t$ should be the same as that of $s^2(l)$ and $S_t$ from the model with short memory error and no intercept. Hence the followings are immediate:

$$\begin{aligned}\widetilde{s}^2(l)\xrightarrow{p}\sigma^2,\ \frac{1}{\sqrt{T}}\widetilde{S}_{[rT]} &\Rightarrow \sigma W(r), \quad (2.45)\\ \frac{1}{T^2}\sum_{t=2}^{T}\widetilde{S}_t^2 &\Rightarrow \sigma^2\int_0^1 W(r)^2dr, \text{ and}\\ \widehat{\eta}_\mu^d &\Rightarrow \int_0^1 W(r)^2dr.\end{aligned}$$

Note that the weak limit of $\widehat{\eta}_\mu^d$ is a functional of a standard Wiener process instead of a Brownian bridge process (for the KPSS test of short memory) or a demeaned Brownian motion (for the KPSS unit root test). This is because we difference the data instead of demeaning the terms in $\widetilde{S}_t$.

**Proof of Theorem 9**

Rewrite the numerator as $T^{-2}\sum_{t=2}^{T}\widetilde{S}_t^2 = T^{-2}\sum_{t=2}^{T}(\epsilon_t - \epsilon_1)^2 = T^{-2}\sum_{t=2}^{T}\epsilon_t^2 - 2\epsilon_1 \cdot T^{-2}\sum_{t=2}^{T}\epsilon_t + T^{-1}\epsilon_1^2$. Multiplying by $T$ gives

$$T^{-1}\sum_{t=2}^{T}\widetilde{S}_t^2 = \frac{1}{T}\sum_{t=2}^{T}\epsilon_t^2 - 2\epsilon_1\frac{1}{T}\sum_{t=2}^{T}\epsilon_t + \epsilon_1^2 \tag{2.46}$$

$$= \frac{1}{T}\sum_{t=2}^{T}\epsilon_t^2 + \epsilon_1^2 + o_p(1),$$

since $\frac{1}{T}\sum_{t=2}^{T}\epsilon_t \xrightarrow{p} 0$. Therefore,

$$T^{-1}\sum_{t=2}^{T}\widetilde{S}_t^2 \xrightarrow{d} \gamma_0 + \epsilon_1^2. \tag{2.47}$$

Second, to figure out the limiting behavior of $\widetilde{s}^2(l)$, rewrite $\widetilde{\gamma}_s$ as below.

$$\widetilde{\gamma}_s = \frac{1}{T}\sum_{t=s+2}^{T}(\epsilon_t - \epsilon_{t-1})(\epsilon_{t-s} - \epsilon_{t-s-1})$$

$$= \frac{1}{T}\left(\sum_{t=s+2}^{T}\epsilon_t\epsilon_{t-s} - \sum_{t=s+2}^{T}\epsilon_t\epsilon_{t-s-1} - \sum_{t=s+2}^{T}\epsilon_{t-1}\epsilon_{t-s} + \sum_{t=s+2}^{T}\epsilon_{t-1}\epsilon_{t-s-1}\right).$$

Plugging this into $\widetilde{s}^2(l)$ yields:

$$\widetilde{s}^2(l) = \left(\frac{1}{T}\sum_{t=2}^{T}\epsilon_t\epsilon_t + 2\sum_{s=1}^{l}w(s,l)\frac{1}{T}\sum_{t=s+2}^{T}\epsilon_t\epsilon_{t-s}\right) \tag{2.48}$$

$$-\left(\frac{1}{T}\sum_{t=2}^{T}\epsilon_t\epsilon_{t-1} + 2\sum_{s=1}^{l}w(s,l)\frac{1}{T}\sum_{t=s+2}^{T}\epsilon_t\epsilon_{t-s-1}\right)$$

$$-\left(\frac{1}{T}\sum_{t=2}^{T}\epsilon_{t-1}\epsilon_t + 2\sum_{s=1}^{l}w(s,l)\frac{1}{T}\sum_{t=s+2}^{T}\epsilon_{t-1}\epsilon_{t-s}\right)$$

$$+\left(\frac{1}{T}\sum_{t=2}^{T}\epsilon_{t-1}\epsilon_{t-1}+2\sum_{s=1}^{l}w(s,l)\frac{1}{T}\sum_{t=s+2}^{T}\epsilon_{t-1}\epsilon_{t-s-1}\right).$$

Then the equation (2.48) can be rewritten by collecting the terms according to the time lags of cross products of $\epsilon_m\epsilon_n$'s (i.e. the value of $m-n$).

$$\begin{aligned}\widehat{s}^2(l) &= \frac{1}{T}\sum_{t=2}^{T-1}\left[2-2w(1,l)\right]\epsilon_t\epsilon_t+\frac{\epsilon_T\epsilon_T+\epsilon_1\epsilon_1}{T} \\ &+\frac{1}{T}\sum_{t=3}^{T-1}2\left[2w(1,l)-w(0,l)-w(2,l)\right]\epsilon_t\epsilon_{t-1} \\ &+\frac{2\left[w(1,l)-w(0,l)\right]\cdot\epsilon_T\epsilon_{T-1}+2\left[w(1,l)-1\right]\epsilon_2\epsilon_1}{T} \\ &+\frac{1}{T}\sum_{t=4}^{T-1}2\left[2w(2,l)-w(1,l)-w(3,l)\right]\epsilon_t\epsilon_{t-2} \\ &+\frac{2\left[w(2,l)-w(1,l)\right]\cdot\left[\epsilon_T\epsilon_{T-2}+\epsilon_3\epsilon_1\right]}{T}+\cdots \\ &+\frac{1}{T}\sum_{t=l+1}^{T-1}2\left[2w(l-1,l)-w(l-2,l)-w(l,l)\right]\epsilon_t\epsilon_{t-l+1} \\ &+\frac{2\left[w(l-1,l)-w(l-2,l)\right]\cdot\left[\epsilon_T\epsilon_{T-l+1}+\epsilon_l\epsilon_1\right]}{T} \\ &+\frac{1}{T}\sum_{t=l+2}^{T-1}2\left[2w(l,l)-w(l-1,l)\right]\epsilon_t\epsilon_{t-l} \\ &+\frac{2\left[w(l,l)-w(l-1,l)\right]\cdot\left[\epsilon_T\epsilon_{T-l}+\epsilon_{l+1}\epsilon_1\right]}{T}-\frac{1}{T}\sum_{t=l+2}^{T}2w(l,l)\epsilon_t\epsilon_{t-l-1}.\end{aligned} \tag{2.49}$$

Now, fix $l$ and let $T$ increase to infinity. Because $\max_{1\le s,t\le T}|\epsilon_t\epsilon_s| = O_p(T)$ [4] one can obtain the following as $T$ increases:

$$\widehat{s}^2(l)\xrightarrow{p}\left[2-2w(1,l)\right]\gamma_0+2\left[2w(1,l)-w(0,l)-w(2,l)\right]\gamma_1 \tag{2.50}$$

[4] Notice that $\max_{1\le s,t\le T}|\epsilon_t\epsilon_s| = \max_{1\le t\le T}\epsilon_t^2$ and recall that for $\epsilon_t\sim I(1)$, $\max_{1\le t\le T}\left(\frac{\epsilon_t}{\sqrt{T}}\right)^2 = O_p(1)$, or equivalently $\max_{1\le t\le T}\epsilon_t^2 = O_p(T)$. Hence we conclude that $\max_{1\le t\le T}\epsilon_t^2 = O_p(T)$ when $\epsilon_t$ is a stationary short or long memory process.

$$+\cdots+2\left[2w(l-1,l)-w(l-2,l)-w(l,l)\right]\gamma_{l-1}+2\left[2w(l,l)-w(l-1,l)\right]\gamma_l-2w(l,l)\gamma_{l+1}.$$

Note that with the Bartlett kernel $w(j,l)=1-\frac{j}{l+1}$, (2.49) can be simplified because

$$2w(j,l)-w(j-1,l)-w(j+1,l)=0$$

for $j=1,2,...,l$. Hence

$$\widetilde{s}^2(l)\xrightarrow{p}\left[2-2w(1,l)\right]\gamma_0-2w(l,l)\gamma_{l+1}=\frac{2}{l+1}\left(\gamma_0-\gamma_{l+1}\right). \tag{2.51}$$

Therefore, $l\cdot\widetilde{s}^2(l)\xrightarrow{p}2\gamma_0$ as $l$ increases since $\gamma_{l+1}\rightarrow 0$ under the assumption of either stationary short- or stationary long-memory process.

Now it is straightforward to see that

$$\frac{T}{l}\widehat{\eta}_\mu^d=\frac{T}{l}\left(\frac{T^{-2}\sum_{t=2}^T\widetilde{S}_t^2}{\widetilde{s}^2(l)}\right)=\frac{T^{-1}\sum_{t=2}^T\widetilde{S}_t^2}{l\widetilde{s}^2(l)}, \tag{2.52}$$

and therefore, using (2.47) and (2.51),

$$\frac{T}{l}\widehat{\eta}_\mu^d\xrightarrow{d}\frac{\gamma_0+\epsilon_1^2}{2\gamma_0}. \tag{2.53}$$

This implies that $\widehat{\eta}_\mu^d\xrightarrow{p}0$ as $l\rightarrow\infty$, $T\rightarrow\infty$, $\frac{l}{T}\rightarrow 0$.

**Proof of Proposition 9**

In this case, $\Delta\epsilon_t\sim I(d_*)$ where $d_*=d-1$ with $-\frac{1}{2}<d^*<0$. This means $\Delta\epsilon_t$ is an anti-persistent process. From Table 3.1 in Section 3.3.1,

$$\begin{aligned}\frac{1}{T^{2(d_*+1)}}\widetilde{S}_{[rT]} &\equiv \frac{1}{T^{d_*+1/2}}\sum_{t=2}^{[rT]}\Delta\epsilon_t\\ &= \frac{1}{T^{d_*+1/2}}\sum_{t=1}^{[rT]}\Delta\epsilon_t-\frac{1}{T^{d_*+1/2}}\Delta\epsilon_1\\ &\Rightarrow \omega_{d_*}W_{d_*}(r),\end{aligned}$$

so

$$\frac{1}{T^{2(d_*+1)}}\sum_{t=2}^{T}\widehat{S}_t^2 \Rightarrow \omega_{d_*}^2\int_0^1 W_{d_*}(r)^2dr,$$

and

$$l^{-2d_*}\widehat{s}^2(l)\xrightarrow{p}\omega_{d_*}^2.$$

Therefore,

$$\left(\frac{l}{T}\right)^{2d_*}\widehat{\eta}_\mu^d \Rightarrow \int_0^1 W_{d_*}(r)^2dr.$$

That is, $\widehat{\eta}_\mu^d = O_p\left(\left(\frac{T}{l}\right)^{2d_*}\right)$ and $\widehat{\eta}_\mu^d$ goes to zero as $l\to\infty$, $T\to\infty$, $\frac{l}{T}\to 0$.

**Proof of Theorem 10**

Since $\epsilon_t \sim I(1/2)$, it it true that $\Delta y_t = \Delta\epsilon_t \sim I(-1/2)$. Fix $l$ and increase $T$ to get $\widehat{s}^2(l)\to\sigma^2(l) = \gamma_0^* + 2\sum_{s=1}^{l} W_{s,l}\gamma_s^*$, where $\gamma_s^* = E\left(\Delta\epsilon_t\Delta\epsilon_{t-s}\right)$. As in Lee and Schmidt (1996, p.291), one can show

$$(l+1)\sigma^2(l) = (l+1)\,\gamma_0^* + 2\sum_{s=1}^{l}(l+1-s)\,\gamma_s^* \tag{2.54}$$

$$= var\left(\sum_{j=2}^{l+2}\Delta\epsilon_j\right) = var\left(\epsilon_{l+2}-\epsilon_1\right)$$

$$= var(\epsilon_{l+2}) + var(\epsilon_1) - 2\rho\sqrt{var(\epsilon_{l+2})}\sqrt{var(\epsilon_1)},$$

where $\rho$ is the correlation between $\epsilon_{l+2}$ and $\epsilon_1$. Recall from equation (2.38) that

$$var(\epsilon_t) = \frac{2\sigma_u^2}{\pi}L(t),$$

where $L(t) = 4\sum_{j=1}^{t}\frac{1}{2j-1}$ and $\frac{L(t)}{\ln t}\to 1$ as $t$ increases. Divide equation (2.54) by $\ln(l+1)$ and let $l$ increase to yield

$$\frac{(l+1)\sigma^2(l)}{\ln(l+1)} \to \frac{2\sigma_u^2}{\pi} \text{ as } l \text{ grows.}$$

This is due to the facts that $var(\epsilon_1)/\ln(l+1) \to 0$ as $l \to \infty$ and $|\rho| \leq 1$. Hence

$$p\lim \frac{l \cdot \widetilde{s}^2(l)}{\ln l} = p\lim \frac{(l+1)\widetilde{s}^2(l)}{\ln(l+1)} = \frac{2\sigma_u^2}{\pi} \equiv K,$$

which implies $\widetilde{s}^2(l) = O_p\left(\frac{\ln l}{l}\right)$ when $T$ and $l$ grow but $l/T$ goes to zero.

Next, consider the sum of $\widetilde{S}_t^2$ in the numerator of $\widehat{\eta}_\mu^d$. Look at the absolute value of the appropriately scaled sum and see

$$\left|\frac{1}{T\ln T}\sum_{t=2}^{T}\widetilde{S}_t^2\right| = \left|\frac{1}{T}\sum_{t=2}^{T}\frac{1}{\sqrt{\ln T}}(\epsilon_t - \epsilon_1)\cdot\frac{1}{\sqrt{\ln T}}(\epsilon_t - \epsilon_1)\right|$$

$$\leq 4\cdot\max_{1\leq t\leq T}\left|\frac{1}{\sqrt{\ln T}}\epsilon_t\right|^2 \simeq 4\cdot\max_{1\leq t\leq T}\left|\frac{K^{1/2}}{s_T}\cdot\epsilon_t\right| \text{ with large } T$$

$$= 4K^{1/2}\cdot\max_{0\leq r\leq 1}|X_T(r)| \Rightarrow 4K^{1/2}\cdot\max_{0\leq r\leq 1}|W(r)| = O_p(1).$$

Therefore, one can obtain

$$\widehat{\eta}_\mu^d = \frac{T^{-2}\sum_{t=2}^{T}\widetilde{S}_t^2}{\widetilde{s}^2(l)} = \frac{\frac{l}{\ln l}\frac{\ln T}{T}\left(\frac{1}{T\ln T}\sum_{t=2}^{T}\widetilde{S}_t^2\right)}{\frac{l}{\ln l}\widetilde{s}^2(l)}$$

$$= \frac{\frac{l}{\ln l}\frac{\ln T}{T}\cdot O_p(1)}{\frac{2\sigma_u^2}{\pi} + o_p(1)} = o_p(1).$$

**Proof of Proposition 12**

Unlike the case for $\widehat{\eta}_\mu$, the calculation of $\widehat{\eta}_\mu^d$ does not involve demeaning so that the full sum of the data $\widetilde{S}_T$ is not zero. So, the correct representation of the HAC estimator with the Bartlett kernel is (see Hashimzade and Vogelsang (2008), page 161)

$$\widetilde{s}^2(l') = \frac{2}{b'(T-1)^2}\sum_{t=2}^{T-1}\widetilde{S}_t^2 - \frac{2}{b'(T-1)^2}\sum_{t=2}^{T-b'(T-1)-1}\widetilde{S}_t\widetilde{S}_{t+b'(T-1)} \tag{2.55}$$

$$-\frac{2}{b'(T-1)^2}\sum_{t=T-b'(T-1)}^{T-1}\widetilde{S}_t\widetilde{S}_T + \frac{1}{T-1}\widetilde{S}_T^2,$$

with $b' = \frac{l+1}{T-1}$.

Now suppose $\epsilon_t$ follows a short-memory process. Plugging $\widetilde{S}_t = \epsilon_t - \epsilon_1$ in (2.55) yields

$$\widehat{s}^2(l') = \frac{2}{b'(T-1)^2}\left(\sum_{t=2}^{T-1}(\epsilon_t-\epsilon_1)^2 - \sum_{t=2}^{T-b'(T-1)-1}(\epsilon_t-\epsilon_1)\left(\epsilon_{t+b'(T-1)}-\epsilon_1\right)\right) \tag{2.56}$$

$$-\frac{2}{b'(T-1)^2}\left(\sum_{t=T-b'(T-1)}^{T-1}(\epsilon_t-\epsilon_1)(\epsilon_T-\epsilon_1)\right)+\frac{1}{T-1}(\epsilon_T-\epsilon_1)^2$$

$$= \frac{2}{b'(T-1)^2}\left(\sum_{t=2}^{T-1}\epsilon_t^2 - 2\epsilon_1\sum_{t=2}^{T-1}\epsilon_t + (T-2)\epsilon_1^2\right) - \frac{2}{b'(T-1)^2}\times$$

$$\left(\sum_{t=2}^{T-b'(T-1)-1}\epsilon_t\epsilon_{t+b'(T-1)} - \epsilon_1\sum_{t=2}^{T-b'(T-1)-1}\left(\epsilon_t+\epsilon_{t+b'(T-1)}\right)+\left(T-b'(T-1)-1\right)\cdot\epsilon_1^2\right)$$

$$-\frac{2}{b'(T-1)^2}\left(\sum_{t=T-b'(T-1)}^{T-1}\epsilon_t\epsilon_T - \sum_{t=T-b'(T-1)}^{T-1}\epsilon_1(\epsilon_t+\epsilon_T)+b'(T-1)\epsilon_1^2\right)+\frac{1}{T-1}(\epsilon_T-\epsilon_1)^2$$

$$\equiv H(b') + \frac{1}{T-1}(\epsilon_T-\epsilon_1)^2.$$

Combining this with $\frac{1}{T^2}\sum_{t=2}^{T}\widetilde{S}_t = \frac{1}{T^2}\left(\sum_{t=2}^{T}\epsilon_t^2 - 2\epsilon_1\sum_{t=2}^{T}\epsilon_t + (T-1)\epsilon_1^2\right)$ gives

$$\widehat{\eta}_\mu^d = \frac{\frac{1}{T^2}\sum_{t=2}^{T}\widetilde{S}_t}{\widehat{s}^2(l')} = \frac{\frac{1}{T}\left(\sum_{t=2}^{T}\epsilon_t^2 - 2\epsilon_1\sum_{t=2}^{T}\epsilon_t + (T-1)\epsilon_1^2\right)}{T\cdot H(b') + \frac{T}{T-1}(\epsilon_T-\epsilon_1)^2}.$$

Denote the weak limit of $\epsilon_T$ as $T$ grows as $\epsilon_\infty$ and apply the functional central limit theorem and continuous mapping theorem to obtain

$$\begin{aligned}\widehat{\eta}_\mu^d &\Rightarrow \frac{\gamma_0+\epsilon_1^2}{\frac{2}{b}(\gamma_0+b'\epsilon_1\epsilon_\infty)+(\epsilon_\infty-\epsilon_1)^2}\\ &= \frac{\gamma_0+\epsilon_1^2}{\frac{2}{b'}\gamma_0+\epsilon_1^2+\epsilon_\infty^2} < \frac{\gamma_0+\epsilon_1^2}{2\gamma_0+\epsilon_1^2+\epsilon_\infty^2} < 1\end{aligned}$$

Secondly, suppose that $\epsilon_t$ follows a unit root process. Rearranging equation (2.56) gives

$$\frac{(T-1)^2}{T^2}\widehat{s}^2(l') = \frac{2}{b'T}\sum_{t=2}^{T-1}\left(T^{-1/2}\epsilon_t\right)^2 - \frac{2}{b'T}\sum_{t=2}^{T-b'(T-1)-1} T^{-1/2}\epsilon_t \cdot T^{-1/2}\epsilon_{t+b'(T-1)}$$

$$-\frac{2}{b'T}\sum_{t=T-b'(T-1)}^{T-1} T^{-1/2}\epsilon_t \cdot T^{-1/2}\epsilon_T + \frac{T-1}{T}\left(T^{-1/2}\left(\epsilon_T - \epsilon_1\right)\right)^2 + o_p(1),$$

where the remaining terms are negligible since

$$T^{-1/2}\epsilon_1 = o_p(1),$$

and

$$\frac{1}{T^2}\sum_{t=2}^{T-1}\epsilon_t\epsilon_1 = T^{-1/2}\epsilon_1 \cdot \frac{1}{T}\sum_{t=2}^{T-1} T^{-1/2}\epsilon_t = o_p(1)O_p(1) = o_p(1).$$

Therefore by applying the functional CLT and continuous mapping theorem one can show

$$\widehat{\eta}_\mu^d \Rightarrow \frac{\int_0^1 W(r)^2 dr}{Q_1^d(b')},$$

where $W(r)$ is the standard Wiener process and

$$Q_1^d(b') \equiv \frac{2}{b'}\left[\int_0^1 W(r)^2 dr - \int_0^{1-b'} W(r)W(r+b')dr - \int_{1-b'}^1 W(r)W(1)dr\right] + W(1)^2.$$

Note that this limit does not degenerate to $\frac{1}{2}$ for $b' = 1$. This is in contrast with the fixed-$b$ limit of $\widehat{\eta}_\mu$ in Proposition 11, which is $\frac{1}{2}$ under both $I(0)$ and $I(1)$ DGPs.

# REFERENCES

REFERENCES

Adenstedt, R. K.: (1974), 'On large-sample estimation for the mean of a stationary random sequence'. *The Annals of Statistics* **2**, 1095–1107.

Amsler, C., P. Schmidt, and T. J. Vogelsang: (2009), 'The KPSS test using fixed-*b* critical values: size and power in highly autocorrelated time series'. *Journal of Time Series Econometrics* **1**.

Andrews, D. W. K.: (1991), 'Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation'. *Econometrica* **59**, 817–854.

Breitung, J.: (2002), 'Nonparametric Tests for Unit Roots and Cointegration'. *Journal of Econometrics* **108**, 343–363.

Caner, M. and L. Kilian: (2001), 'Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate'. *Journal of International Money and Finance* **20**, 639–657.

Davydov, Y. A.: (1970), 'The invariance principle for stationary processes'. *Theory of Probability & Its Applications* **15**, 487–498.

De Jong, R. M. and J. Davidson: (2000), 'Consistency of Kernel Estimators of Heteroskedastic and Autocorrelated Covariance Matrices'. *Econometrica* **68**, 407–424.

Dejong, D. N., J. C. Nankervis, N. E. Savin, and C. H. Whiteman: (1992), 'Integration versus Trend Stationarity in Time Series'. *Econometrica* **60**, 423–434.

Diebold, F. X. and G. D. Rudebusch: (1991), 'On the power of Dickey-Fuller tests against fractional alternatives'. *Economics Letters* **35**, 155–160.

Giraitis, L., H. L. Koul, and D. Surgailis: (2012), *Large sample inference for long memory processes*. Imperial College Press London.

Granger, C. W. and R. Joyeux: (1980), 'An introduction to long-memory time series models and fractional differencing'. *Journal of Time Series Analysis* **1**, 15–29.

Hansen, B. E.: (1992), 'Consistent Covariance Matrix Estimation for Dependent Heterogenous Processes'. *Econometrica* **60**, 967–972.

Hashimzade, N. and T. J. Vogelsang: (2008), 'Fixed-*b* Asymptotic Approximation of the Sampling Behavior of Nonparametric Spectral Density Estimators'. *Journal of Time Series Analysis* **29**, 142–162.

Hassler, U. and J. Wolters: (1994), 'On the power of unit root tests against fractional alternatives'. *Economics Letters* **45**, 1–5.

Hosking, J. R.: (1981), 'Fractional differencing'. *Biometrika* **68**, 165–176.

Jansson, M.: (2002), 'Consistent Covariance Estimation for Linear Processes'. *Econometric Theory* **18**, 1449–1459.

Kiefer, N. M. and T. J. Vogelsang: (2002a), 'Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation'. *Econometrica* **70**, 2093–2095.

Kiefer, N. M. and T. J. Vogelsang: (2002b), 'Heteroskedasticity-Autocorrelation Robust Testing Using Bandwidth Equal to Sample Size'. *Econometric Theory* **18**, 1350–1366.

Kiefer, N. M. and T. J. Vogelsang: (2005), 'A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests'. *Econometric Theory* **21**, 1130–1164.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin: (1992), 'Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure are We that Economic Time Series Have a Unit Root'. *Journal of Econometrics* **54**, 154–178.

Lee, D. and P. Schmidt: (1996), 'On the power of the KPSS test of stationarity against fractionally-integrated alternatives'. *Journal of Econometrics* **73**, 285–302.

Lee, H. S. and C. Amsler: (1997), 'Consistency of the KPSS unit root test against fractionally integrated alternative'. *Economics Letters* **55**, 151–160.

Liu, M.: (1998), 'Asymptotics of nonstationary fractional integrated series'. *Econometric Theory* **14**, 641–662.

Lo, A. W.: (1991), 'Long-Term Memory in Stock Market Prices'. *Econometrica* **59**, 1279–1313.

Mandelbrot, B. B. and J. W. Van Ness: (1968), 'Fractional Brownian motions, fractional noises and applications'. *SIAM Review* **10**, 422–437.

Müller, U. K.: (2005), 'Size and power of tests of stationarity in highly autocorrelated time series'. *Journal of Econometrics* **128**, 195–213.

Newey, W. K. and K. D. West: (1987), 'A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix'. *Econometrica* **55**, 703–708.

Phillips, P. C. B. and P. Perron: (1988), 'Testing for a Unit Root in Time Series Regression'. *Biometrika* **75**, 335–346.

Qiu, J. and Z. Lin: (2011), 'The invariance principle for fractionally integrated processes with strong near-epoch dependent innovations'. *Science China Mathematics* **54**, 117–132.

Said, S. E. and D. A. Dickey: (1984), 'Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order'. *Biometrika* **71**, 599–607.

Shin, Y. and P. Schmidt: (1992), 'The KPSS stationarity test as a unit root test'. *Economics Letters* **38**, 387–392.

Sowell, F.: (1990), ‘The fractional unit root distribution’. *Econometrica* **58**, 495–505.

Tanaka, K.: (1999), ‘The nonstationary fractional unit root’. *Econometric Theory* **15**, 549–582.

Taqqu, M. S.: (1975), ‘Weak convergence to fractional Brownian motion and to the Rosenblatt process’. *Probability Theory and Related Fields* **31**, 287–302.

Vogelsang, T. J. and M. Wagner: (2013), ‘A fixed-b perspective on the Phillips–Perron unit root tests’. *Econometric Theory* **29**(03), 609–628.